PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07H 19/06, 19/16, C08L 77/00
A61K 31/785, 31/505, 31/52

A1

(11) International Publication Number: WO 93/12129

(43) International Publication Date: 24 June 1993 (24.06.93)

(21) International Application Number: PCT/US92/10921

(22) International Filing Date: 17 December 1992 (17.12.92)

(30) Priority data:
07/809,661 18 December 1991 (18.12.91) US

(71) Applicant *(for all designated States except US)*: GLAXO INC. [US/US]; Five Moore Drive, Research Triangle Park, NC 27709 (US).

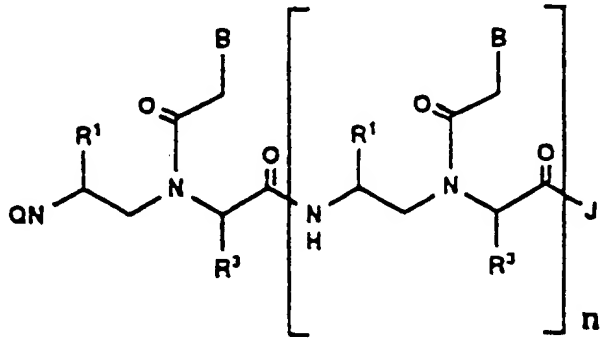
(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : THOMSON, Stephen, A. [US/US]; 106 Winterberry Ridge Drive, Durham, NC 27713 (US). NOBLE, Stewart, A. [GB/US]; 4412 Turnberry Circle, Durham, NC 27712 (US). RICCA, Daniel, J. [US/US]; 5855 Kiger Road, Rougemont, NC 27572 (US).

(74) Agents: LEVY, David, J. et al.; Glaxo Inc., Five Moore Drive, Research Triangle Park, NC 27709 (US).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).

Published
With international search report.

(54) Title: PEPTIDE NUCLEIC ACIDS AND THEIR EFFECT ON GENETIC MATERIAL



(I)

(57) Abstract

Peptide nucleic acids oligomers of formula (I) wherein n is 1 or more, particularly about 5-20 and B is independently one of the 4 nucleoside bases or their equivalents, Q and J are end groups useful in anti-sense oligomers and their use in affecting genetic material, e.g. as triplex or antisense in the treatment of disease.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

Peptide Nucleic Acids and Their Effect on Genetic Material

Background of the Invention

The prevention of gene transcription and/or gene translation at the DNA/mRNA level is attractive for many reasons. Classical approaches to drug discovery involve the design and identification of compounds directed against unrelated proteins such as enzymes, receptors or ion channels, the structure and mode of action of which are usually very complicated and often poorly understood. Conversely, the potential for therapeutic intervention at the nucleic acid level follows a well ordered, generalizable strategy which is targeted at the initiating events of an amplifying cascade; thus transcription of a gene gives rise to many copies of mRNA which on translation affords an even greater number of protein molecules. Inhibition of gene expression ought, therefore, to be more efficient than inhibition of the gene product.

Anticancer therapy using DNA binding or modifying drugs is well established, however, current agents such as doxorubicin, mitoxantrone and cisplatin (Scrip's Cancer Chemotherapy Review 1991) are not capable of recognizing specific gene sequences, and therefore, lack selectivity, discriminating only poorly between cancer and normal cells. A synthetic oligodeoxynucleotide (ODN), can provide absolute specificity of action since statistically the sequence defined by any linear combination of the four heterocyclic bases, adenine (A), guanine (G), cytosine (C), and thymine (T), to form an oligonucleotide of 17 residues in length, occurs just once in the entire sequence of the human genome. Thus, the ODN can bind via Watson-Crick or Hoogsteen base pairing to its complementary base sequence which could, for example, be part of an oncogene implicated in tumorigenesis or an element of genetic material implicated as the dominant cause of a disease phenotype, for instance, a sequence which comprises an essential target within a viral genome.

The potential of such 'antisense' (AS) oligodeoxynucleotides to serve as code blocking therapeutic principles was recognized by Zamecnik and Stephenson (Proc. Natl. Acad. Sci. USA, 1978, 75, 280) who demonstrated the inhibition of Rous sarcoma virus replication in chick embryo fibroblasts on addition of a tridecamer ODN with sequence complementarity to reiterative sequences in the viral (+) RNA genome. Duplex formation, i.e. the interaction of an AS ODN with an

RNA molecule, as a means to achieve inhibition of gene function has received considerable attention of late. The precise mechanism by which inhibition occurs is dependent on the micro-environment of the target sequence. Thus an AS ODN directed against a ribosome binding site would be expected to competitively inhibit translation initiation whereas those complementary to a coding sequence might cause hybridization arrest of translation by preventing ribosomal translocation along the RNA (Bacon and Wickstrom, Oncogene Research, 1991, 6, 13). Alternatively, complexation might lead to cleavage of the target RNA via the intermediacy of RNase H (Shuttleworth and Colman, EMBO J., 1988, 7, 427). Interception of post-transcriptional processes such as RNA splicing is also possible by judicious choice of sequence and has proved particularly effective against viral targets, e.g. Herpes simplex Virus (HSV) (Smith et al., Proc. Natl. Acad. Sci. USA, 1986, 83, 2787) and Human Immunodeficiency Virus (HIV) (Goodchild et al., Proc. Natl. Acad. Sci. USA, 1988, 85, 5507), where alternative splicing of precursor RNAs is commonly used as a strategy to achieve control of viral replication.

The therapeutic potential of AS ODNs is the subject of several recent reviews (Zon, Nucleoside Analogs as Antiviral Agents, ACS Symposium Series, Martin, J.C. Ed., American Chemical Society, Washington, DC. 1989, p170; Weintraub, Scientific American, 1990, 40: Uhlmann and Peyman, Chemical Reviews, 1990, 90 (4), 543; Matteucci and Bischofberger, Annual Reports in Medicinal Chemistry, 1991, 26, 287). This therapeutic potental is further highlighted by a recent publication citing the selective action of unmodified AS ODNs targetted to c-myb on leukemic hematopoetic cells (Calabretta et al, Proc. Natl. Acad Sci. USA, 1991, 88, 2351). The authors claim that perturbation of c-myb function with AS ODNs might form the basis of a molecular approach to leukemia therapy, perhaps most immediately as ex vivo bone marrow purging agents. Another paper (Calabretta et al, Proc. Natl. Acad. Sci. USA, 1991, 88) details experiments in which unmodfied AS ODNs were directed against the bcr-abl splice to beneficial effect. A patent by the same authors describe AS ODNs against the c-abl proto-oncogene (WO 91/03260). It is of note that Genta, Inc. of San Diego, California, recently filed an Investigational New Drug (IND) application with the Food and Drug Administration to begin trials of an AS ODN to treat Chronic Myelogenous Leukemia (Sci Com Cancer Report, 1991, 10). Also of note are the applications of unmodified oligomers to address cystic fibrosis (Sorscher et al., Proc. Natl. Acad. Sci, USA, 1991, 88, 7759), Herpes Simplex Virus (HSV) (WO 90/07409) and Human Papillomavirus (HPV) (WO 91/08313).

The inhibition of transcription by direct action on DNA itself, where copy number is restricted to two per cell, is an even more attractive target than inhibition of translation by therapeutic intervention at the RNA level. The sequence specific recognition of double helical DNA by synthetic ligands is the subject of recent reviews (Dervan, Nucleic Acids and Molecular Biology, Vol 2, ed., F. Eckstein and D.M. Lilley, Springer-Verlag, 1988, p49; Nielsen, Bioconjugate Chemistry, 1991, 2 (1), 1). Of particular interest to the current invention is the design of ODNs and their analogues which bind to ds DNA forming triple stranded structures, i.e. a "triplex", using the structural motifs first described by Hoogsteen (Acta. Cryst., 1959, 12, 822). It is recognized that the precise binding motif adopted by the ligand might vary from that cited above as postulated by Birg et al., Nucleic Acidi Research, 1990, 18 (10), 2901. Furthermore interactive or reactive groups might be appended to the ligand to beneficial effect (Shaw et al., J. Amer. Chem. Soc., 1991, 113, 7765). That such an approach has utility with regard to therapeutic intervention is evidenced by several recent publications.

Thus triple helix formation has been shown to inhibit the function of DNA binding proteins (Maher et al., Science 1989, 245, 725, Orson et al., Nucleic Acids Research 1991, 19(12) 3435) and to effect inhibition of transcription elongation in vitro (Young et al., Proc. Natl. Acad. Sci. USA 1991, 88, 10023). Most recently it has been shown that an ODN binds to the promoter region of c-myc in HeLa cells, thereby selectively reducting c-myc RNA levels. The therapeutic potential of this approach has also been highlighted in the recent patent literature (WO 90/15884, EP 0375408).

Most recently reports of analogues containing amide bonds have appeared in the art (Weller et al., J. Org. Chem., 1991, 56, 6000; Huang et al., ibid., 1991, 56, 6007). At the Twelth American Peptide Symposium at the Massachusetts Institute of Technology in Cambridge, Massachusetts on June 17, 1991, Rolf Berg of the RISO National Laboratory in Roiskilde, Denmark presented work on modified peptides with nucleoside side chains which were called peptide nucleic acids (PNAs). However, only PNAs from the T monomer could be made. Presentations by this group on July 8, 1991 at the University of California at Berkeley set forth descriptions of certain PNAs. A publication of their work is by Peter E. Nielsen et al. in Science, Vol. 254, pages 1497-1500 (6 December 1991).

Summary of the Invention

Oligomers having at least one peptide bond in the backbone with at least one pendant purine or pyrimidine nucleoside base are useful in affecting genetic material for diagnostic, therapeutic or analytic purposes.

Brief Description of the Drawings

Fig. 1 depicts a schematic representation of a process used to make a particular peptide nucleic acid (PNA) of the invention.
Fig. 2 is a schematic of a test used to determine the degree of binding of a PNA according to the invention to genetic material.
Fig 3. is a graph showing the variation with increasing PNA concentration of binding to genetic material.

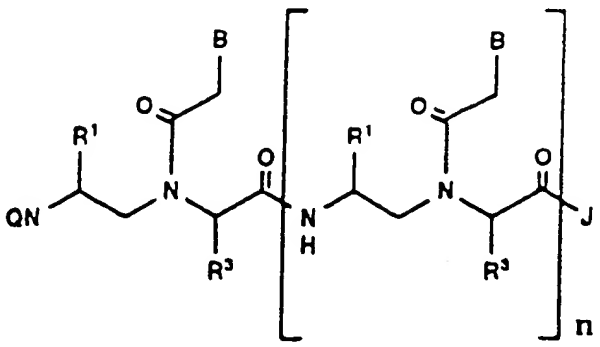
Detailed Description of the Invention

Nucleoside base oligomers which have at least one purine or pyrimidine nucleoside base bound to a backbone having at least one peptide bond constitute the present invention. Preferably, the backbone would have 1 peptide bond for each pendant base whereby the oligomer can be formed from monomers each having an A, T, G or C nucleoside base. By selecting the A,T, G or C amino acid monomers, each amino acid of the oligomer can be built up by successive peptide bond formations.

The particular number of nucleoside bases in a PNA of the invention will depend on the use to which the PNA is put, i.e. the target portion of genetic material. Below 6 nucleoside bases, there will usually be too many possible different targets within the genetic material, e.g. many different chromosomes have a portion with GATT as a subsequence. Above 16 bases, the additional specificity provided is unnecessary, i.e. there will only be 1 sequence with a particular 15 base arrangement and no further purpose is provided by the additional bases.

In addition to the backbone and bases, the peptide oligomers of the invention may have pendant groups, usually at the termini, to stabilize the end, to act as an intercalator, to facilitate cellular uptake or to increase solubility.

A particular peptide oligomer of the invention is that of the following formula (I):



(I)

wherein

Q is an N-terminal blocking group;

J is a C-terminal blocking group or Q and J may together be a single bond;

n is at least 1;

R^1 is independently hydrogen, benzyl, $-CH_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl, $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO(C_{1-4}$ alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkynyl, C_{2-8} alkenyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl)substituted amino;

R^3 is independently hydrogen, benzyl, $-CH_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl, $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO(C_{1-4}$ alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkynyl, C_{2-8} alkenyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl)substituted amino;

B is independently a monovalent purine or pyrimidine nucleoside base,i.e. a base such as guanine without the hydrogen at the 9-position

or an acid- or base- addition salt thereof.

Q is preferably an N-terminal blocking group which may stabilize that portion of the molecule, e.g. sterically hindered alkanoyl group whereby an amide is formed by the group QNH-. Another function of the N-terminal blocking group Q is as an intercalator to bind within the genetic material, e.g. to actually wedge itself within the

DNA double helix as described in Oligodeoxynucleotides-Antisense Inhibitors of Gene Expression, ed. by Jack S. Cohen, MacMillan Press, London (1989) (ISBN 0-333-49211-0). When used to increase cellular upake, Q may function to increase lipophilicity, e.g. a steriodal moiety, or to induce active uptake such as by having a recognition moiety for the cell surface, e.g. a steroidal molety or a saccharide moiety. To increase solubility, Q may contain an ionizable moiety such as a carboxylic acid or an amine, e.g. the QNH molety may be $NH_2(CH_2)_4CH(NH_2)CONH$-.

J may be any of the types of groups described above for Q. Specific examples include, as the -CO-J group, -COOt-butyl as a sterically hindered stabilizing group and $-CONHCH(CONH_2)CH_2CH_2CH_2CH_2NH_2$ as an ionizable moiety which functions as a solubizing moiety.

n is at least 1 with compounds wherein n=1-4 being useful as intermediates for those wherein n is at least about 5, e.g. - about 5 to 20 although sequences where n is between 6 to 16 are particular aspects of the invention.

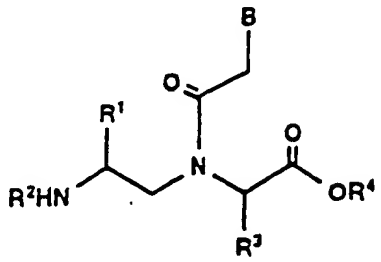
R^1 is preferably hydrogen, benzyl, $-CH_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO(C_{1-4}$ alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkynyl, C_{2-8} alkenyl, e.g. $-CH_2CH=CHCH_3$, e.g. $-(CH_2)_4CCH$, C_{5-8} cycloalkyl, e.g. cyclohexyl aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl)substituted amino. As values for C_{1-4} alkyl in any of such definitions of R^1, e.g. alkoxy, these may be methyl, ethyl, iso-propyl, n-propyl, n-butyl, sec-butyl, iso-butyl and tert-butyl. As values of aryl there are incuded phenyl and naphthyl and for heteroaryl, there are included 5- and 6-membered rings with 1,2 or 3 N,O or S heteroatoms with the proviso that two O or S atoms are not bonded to each other with examples being pyridinyl, oxazolyl, thienyl, thiadiazolyl and triazolyl with attachment through a carbon or nitrogen atom, e.g. $-N(CH=CH)_2$. Halogen includes chloro, bromo, iodo and fluoro.

R^3 is preferably hydrogen, benzyl, $-Ch_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl, $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO(C_{1-4}$ alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8}

alkenyl, C_{2-8} alkenyl, e.g. $-CH_2CH=CHCH_3$, e.g. $-(CH_2)_4CCH$, C_{5-8} cycloalkyl, e.g. cyclopentyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl)substituted amino. As values for C_{1-4} alkyl in any of such definitions of R^1, e.g. alkoxy, these may be methyl, ethyl, iso-propyl, n-propyl, n-butyl, sec-butyl, iso-butyl and tert-butyl. As values of aryl there are included phenyl and naphthyl and for heteroaryl, there are included 5- and 6-membered rings with 1,2 or 3, N,O or S heteroatoms with the proviso that two O or S atoms are not bonded to each other with examples being pyridinyl, oxazolyl, thienyl, thiadiazolyl and triazolyl with attachment through a carbon or nitrogen atom, e.g. $-N(CH=CH)_2$. Halogen includes chloro, bromo, iodo and fluoro.

B is a purine or pyrimidine nucleoside base is preferably adenine, thymine, guanine or cytosine or an equivalent thereof which binds to its complement, i.e. adenine to thymine and guanine to cytosine. Examples of such equivalents are 5-methylcytosine, 5-propynyluracil, 7-propynyl-7-deaza-adenine and 7-methyl-7-deaza-adenine. Preferably, the peptide oligomer of the invention has at least 3 different A, T, G and C bases or their equivalent, e.g. all four of such bases.

Also part of the present invention are monomers of the following formula (X) which may be polymerized to yield the oligomer of formula (I)



(X)

wherein

R^1 is as defined for formula (I);
R^2 is an amino protecting group;
R^3 is as defined for formula (I);
R^4 is a carboxylic acid protecting group;
B is as defined for formula (I),
or an acid-or base-addition salt thereof.

R^2 is preferably t-butyloxycarbonyl, 9-fluorenylmethoxycarbonyl, carbobenzoxy (i.e. benzyloxy carbonyl) trityl or dimethoxytrityl.

R^4 is preferably alkyl, e.g. methyl, ethyl, tert-butyl, or (2-trimethylsilyl)ethyl, aryl, e.g. phenyl or benzyl.

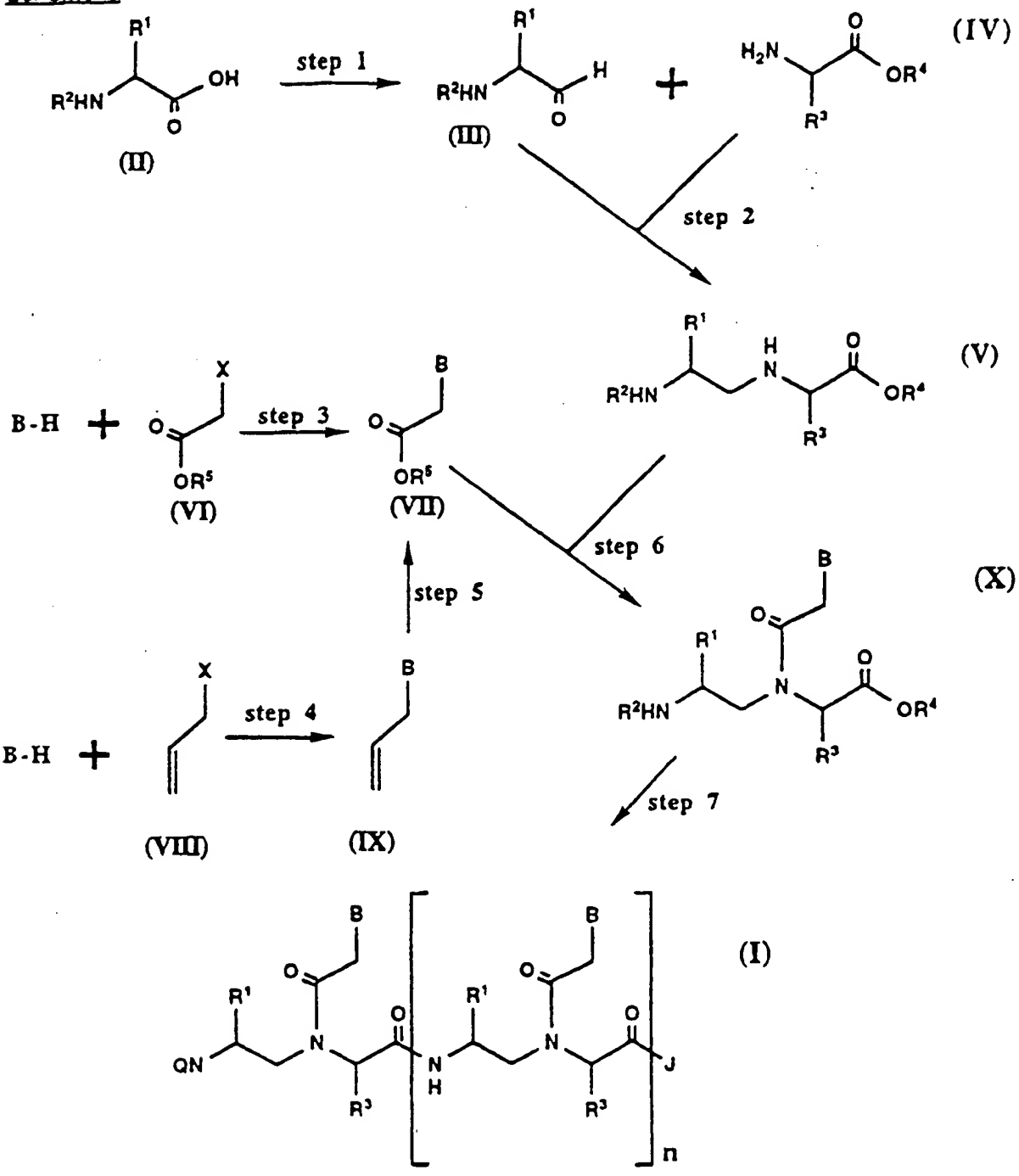
Also part of the present invention are novel intermediates and processes, e.g. the di-, tri- and tetra- peptide oligomers which are used as intermediates to produce the AS oligomers of formula (I).

In the oligomer of formula (I) and the monomer of formula (X), several asymmetric centers are present. The present invention encompasses all isomers and mixtures thereof within the scope of all the formulae provided. For example, the carbon bearing the R^1 and R^3 groups may independently each be R or S to give the isomers RR, RS, SS and SR.

Process

The compounds of formula (I) may be prepared by the pathway outlined in Scheme 1.

Scheme 1



In step 1, an alpha-amino acid of formula (II) or a derivative thereof, wherein R^1 is as defined above for formula (I) and R^2 is defined for formula (X) is reduced by methods known in the literature (see Janusz Jurczak, Chem. Rev. 1989, 89, 149) to yield a compound of formula (III). For example the ethyl ester of the compound of formula (I) is treated with diisobutylaluminum hydride at -78 C to give the compound of formula (III).

R^1 in formula (II) and R^3 in formula (IV) may be used in a protected form to avoid reactivity of these groups during subsequent steps such as steps 1,2,6 and 7. Thus, if R^1 is $CH_2CH_2CH_2CH_2NH_2$ forming a lysine sidechain, the starting material of formula (II) may be $BocNHCH(CH_2CH_2CH_2CH_2NHCOOCH_2C_6H_5)COOH$, wherein the benzyloxycarbonyl group may be removed after preparation of the final compound of formula (I) by treatment with hydrogen fluoride, or hydrogenation with H_2 over a noble metal catalyst.

In step 2, a compound of formula (III) is reacted with a compound of formula (IV) in a reductive amination to yield a compound of formula (V). In the compound of formula (IV), R^3 is as defined above for formula (I) and R^4 is a carboxylic acid protecting group as defined for formula (X) such as alkyl (e.g. methyl). The carboxylic acid protecting group maintains the COO- group of formula (IV) through the reductive amination conditions of step 2 and the amide bond forming conditions of step 6. The reaction of step 2 is carried out in a solvent such as methanol, in the presence of a dehydrating agent, e.g. molecular sieves, and a reducing agent such as sodium cyanoborohydride at about 25 ° C as described by Zydowsky et al in J. Org. Chem. 1988, 53, 5607.

This route to the compounds of formula (V) has the advantage over other possible routes in that it allows for independent selection of R^1 and R^3 and independent control of the stereochemistry at the carbon atoms which bears R^1 and R^3. Since the starting materials for this route to compounds of formula (V) are alpha-amino acids the chiral pool of natural and unnatural alpha-amino acids can be used to produce the oligomers of the invention.

In step 3, a compound B-H in which B is defined as in formula (X) or an appropriately protected derivative thereof, for example N-6-benzyloxycarbonyladenine (Az), is reacted with a compound of formula (VI), wherein X is a leaving group, e.g. bromine and R^5 is hydrogen or a commonly used carboxylic acid protecting group, such as tert-butyl, in a suitable solvent, such as

dimethylformamide, under basic conditions, e.g. potassium carbonate, to yield a compound of formula (VII). The attachment of the compound of formula (VI) to B-H is at the 1 position for T and C and at the 9 position for A and G, and at the corresponding positions when B is a nucleobase analog.

Step 4 depicts where, in certain cases, it is of advantage to use a masked equivalent of a compound of formula (VI) such as a compound of formula (VIII) wherein X is a leaving group such as bromine. In step 4, B-H is reacted with 3-bromopropene (formula (VIII) X = Br) to give a compound of formula (IX), e.g. at 27°C in dimethylformamide under basic conditions e.g. potassium carbonate. Step 5 shows the conversion of (IX) to a compound of formula (VII) by oxidative cleavage of the double bond, for example by treatment with sodium periodate in the presence of ruthenium tetraoxide at about 25 ° C as described by Carlsen et al in J. Org. Chem. 1981, 46, 3936.

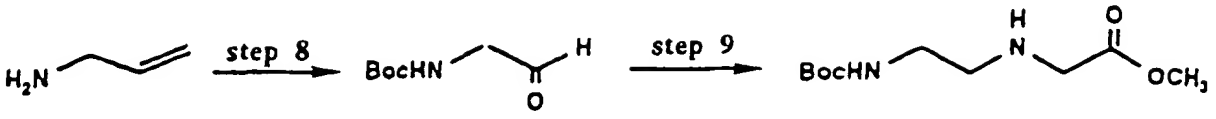
In step 6, a compound of formula (VII) in which R^5 is H is reacted with a compound of formula (V) under conditions known in the art for forming amide bonds to yield a compound of formula (X) (see Miklos Bodanszky; Peptide Chemistry, A Practical Textbook, Springer-Verlag 1988). This may involve conversion of the carboxyl moiety of a compound of formula (VII) to an activated form such as an activated ester, acid chloride, or mixed anhydride, and reaction of this activated form with a compound of formula (V) to give a compound of formula (X). For example, a compound of formula (VII) in which R^5 is hydrogen is activated with benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP), 1-hydroxybenzotriazole (HOBt), in dimethylformamide in the presence of diisopropylethylamine, is reacted with a compound of formula (V) in which R^2 is Boc, and R^4 is methyl to yield a compound of formula (X) in which R^2 is Boc, R^4 is methyl.

In step 7, compounds of formula (X) can be converted to a compound of formula (I) by reacting a compound of formula (X) in which R^2 is hydrogen with a compound of formula (X) in which R^4 is hydrogen under conditions known in the art for forming amide bonds (cf. Miklos Bodanszky; Peptide Chemistry, A Practical Textbook, Springer-Verlag 1988). This may involve conversion of the carboxyl moiety of a compound of formula (X) where R^4 is hydrogen to an activated form such as an activated ester, acid chloride, or mixed anhydride, and reacting this activated form with a compound of formula (X) where R^2 is hydrogen. This coupling reaction can be repeated with monomers of Formula (X) with different B groups to give oligomers

and polymers of a compound of formula (I). The reaction of step 7 can be done using standard solution phase reaction conditions, for example a compound of formula (X) in which R^4 is hydrogen, and R^2 is Boc is reacted with a compound of formula (X) in which R^2 is hydrogen, and R^4 is methyl, in dimethylformamide in the presence of the coupling reagents O-benzotriazol-1-yl-N,N,N',N',-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt), and diisopropylethylamine to yield a compound of formula (I). The coupling can also be performed by anchoring one of the reaction components on a solid support, such as a polystyrene resin and then performing a repetitive cycle of coupling and deprotection steps which allows for the rapid preparation of compounds of formula (I) in which n is greater than 1. This method is commonly known as solid phase synthesis (see Merrifield, J. Am. Chem. Soc. 1963, 85, 2149, and Science 1986, 232, 341). For example, a compound of formula (X) in which R^4 is hydrogen, and R^2 is Boc is coupled to a MBHA resin to which is anchored a lysine (with the epsilon amino group protected) through the carboxyl group in dimethylformamide in the presence of the coupling reagents HBTU, HOBt, and diisopropylethylamine. After the coupling is complete the Boc group is removed by strong acid which reveals a free amino group to which a second residue can be coupled. Repeating this coupling-deprotection cycle five more times and cleaving the chain from the solid support with hydrogen fluoride yields a compound of formula (I) in which n is five, J is lysine, and Q is hydrogen. In many cases some of the functional groups on the bases will be protected to avoid undesired side reactions during the synthesis of the compounds of formula (I).

Protecting groups on the nucleobases must be removed so that they will be able to bind to the target genetic material. The protecting groups can be removed by methods such as treatment with fluoride ion, hydrofluoric acid, or by hydrogenation with H_2 in the presence of a noble metal catalyst. This deprotection can be preformed with the chain attached to the solid support, or after it has been removed.

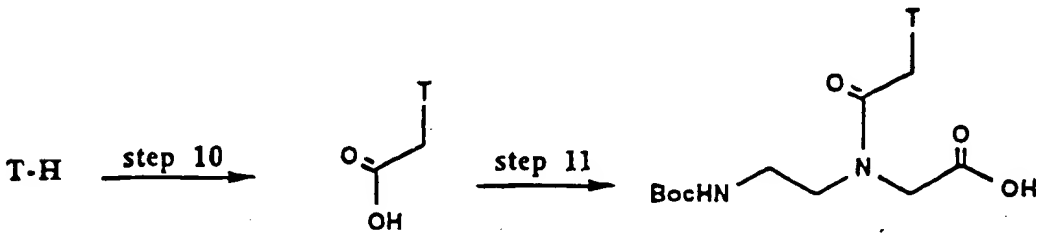
Scheme 2



(XI) (IIIa) (Va)

Scheme 2 depicts a method to make the compound of formula (V) in which R^1 and R^3 are hydrogen, R^2 is Boc, and R^4 is methyl (formula (Va)). In step 8 of this Scheme allyl amine (formula (XI)) is converted to the aldehyde of formula (III) where R^1=H and R^2=Boc (formula (IIIa)), N-tert-butyloxycarbonylglycinal, as described by S.A.Thompson et al in J. Med. Chem. 1986, 29, 104. In step 9 the N-tert-butyloxycarbonylglycinal is reacted with glycine methyl ester hydrochloride, in the presence of sodium acetate, 4 A molecular sieves, and sodium cyanoborohydride in methanol to give the compound of formula (Va).

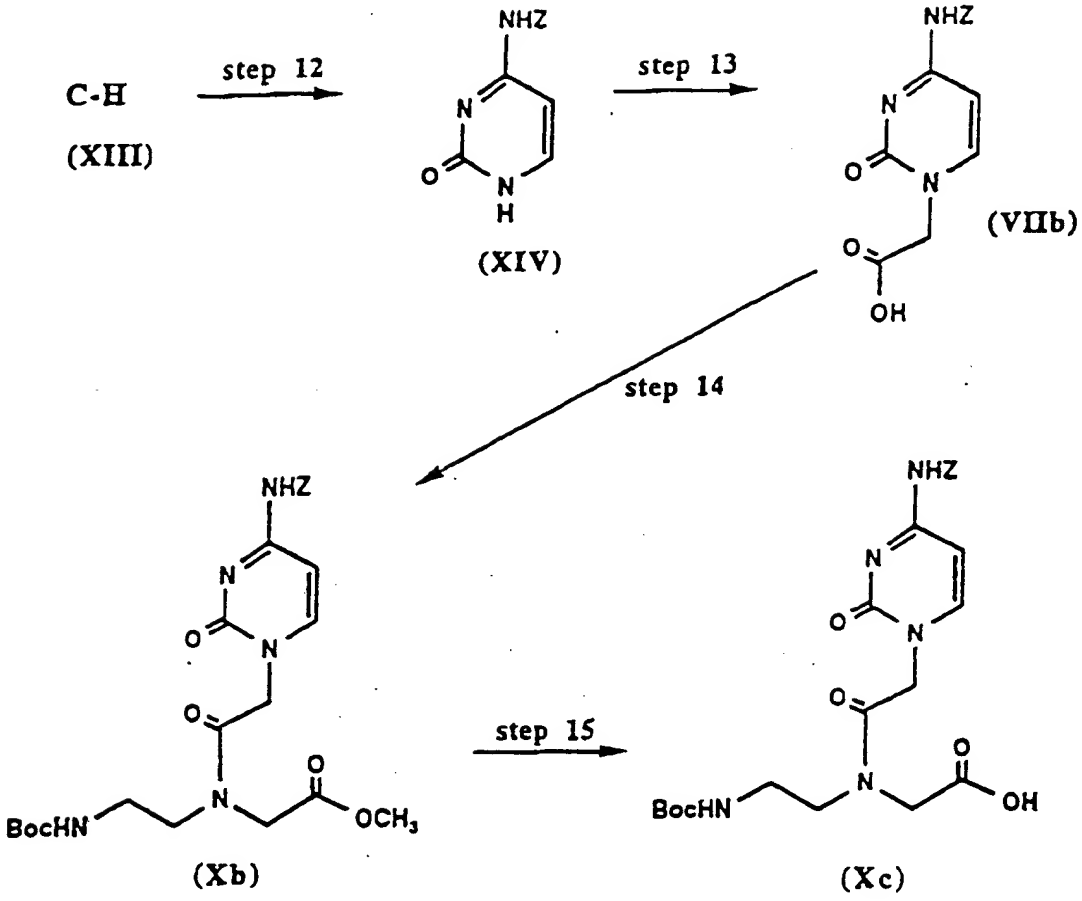
Scheme 3



(XII) (VIIa) (Xa)

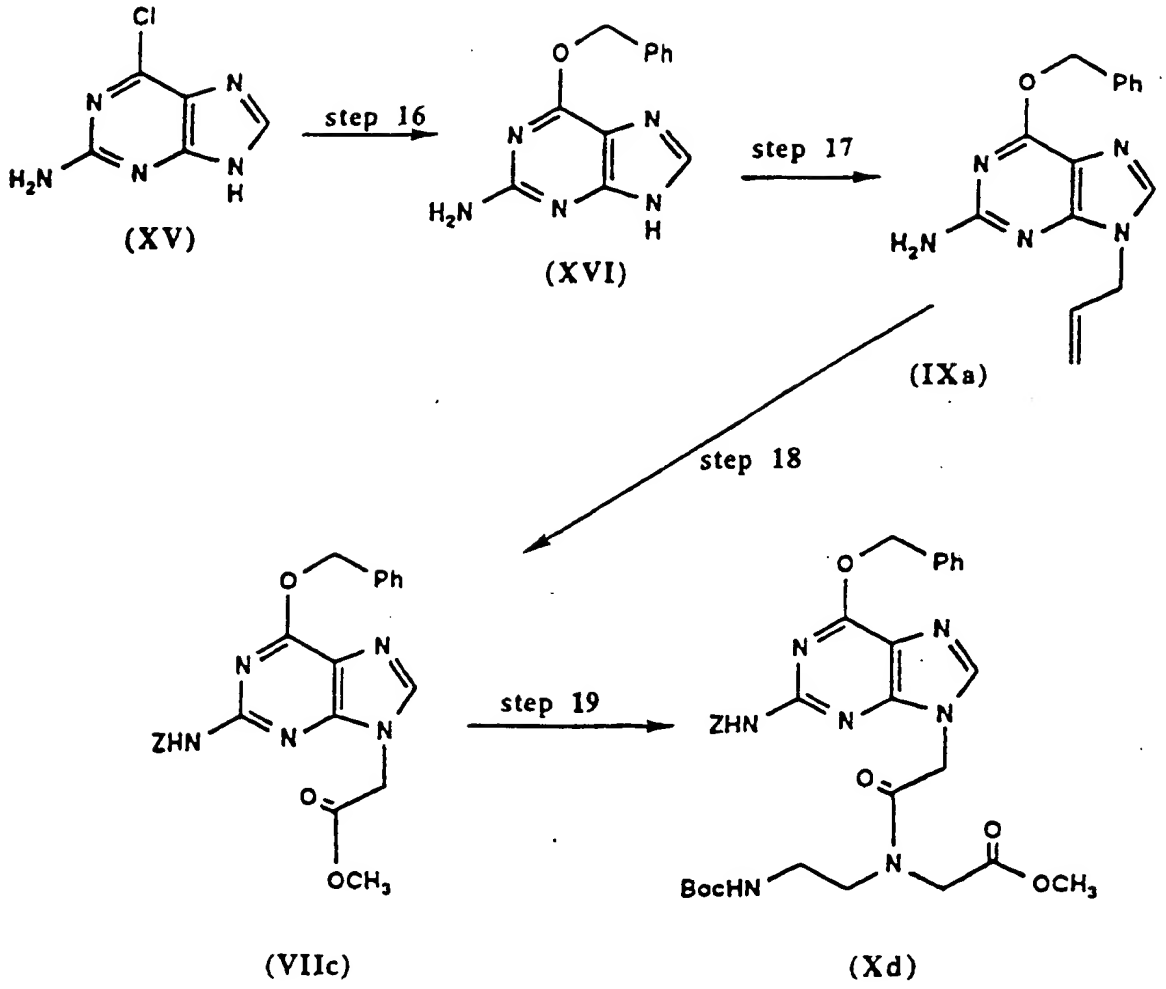
Scheme 3 is a more detailed description of steps 3 and 6 of Scheme 1, and depicts a method for making the compound of formula (X) in which R^1, R^3 and R^4 are hydrogen, R^2 is Boc, and B is thymine (formula (Xa)). In step 10 of this scheme thymine (T-H, Formula (XII)) is reacted with chloroacetic acid in aqueous potassium hydroxide as described by A. S. Jones et al, in Tetrahedron, 1973, 29, 2293, to give the compound of formula (VII) where B=T and R^5=H, formula (VIIa). In step 11 the compound of formula (VIIa) is activated with BOP in dimethylformamide and reacted with the compound of formula (Va), followed by hydrolysis of the resulting methyl ester by treatment with aqueous lithium hydroxide to give the compound of formula (Xa). The compound of formula (Xa) is referred to as the Teg monomer.

Scheme 4



Scheme 4 depicts a synthesis of the monomer of formula (X) in which R^1, R^3 and R^4 are hydrogen, R^2 is Boc, and the nucleobase, B, is 4-N-benzyloxycarbonylcytosine (formula (Xc)). In step 12 the exocyclic amino group of cytosine, formula (XIII), is protected with the benzyloxycarbonyl group (Z) to give the compound of formula (XIV). In step 13 the compound of formula (XIV) is reacted with tert-butyl bromoacetate, which is followed by removal of the t-butyl group with strong acid (trifluoroacetic acid in methylene chloride) to give the compound of formula (VII) where R^5=H and B=protected cytosine, formula (VIIb). In step 14 the compound of formula (VII) where R^5=H and B=protected cytosine, formula (VIIb) is activated with BOP in dimethylformamide and reacted with the compound of formula (Va) to give the compound of formula (X) where R^1=R^3=H, (Xb). In step 15 the methyl ester of the compound of formula (Xb) is hydrolyzed by treatment with aqueous lithium hydroxide to give the compound of formula (Xc), which is referred to as the Z protected Ceg monomer.

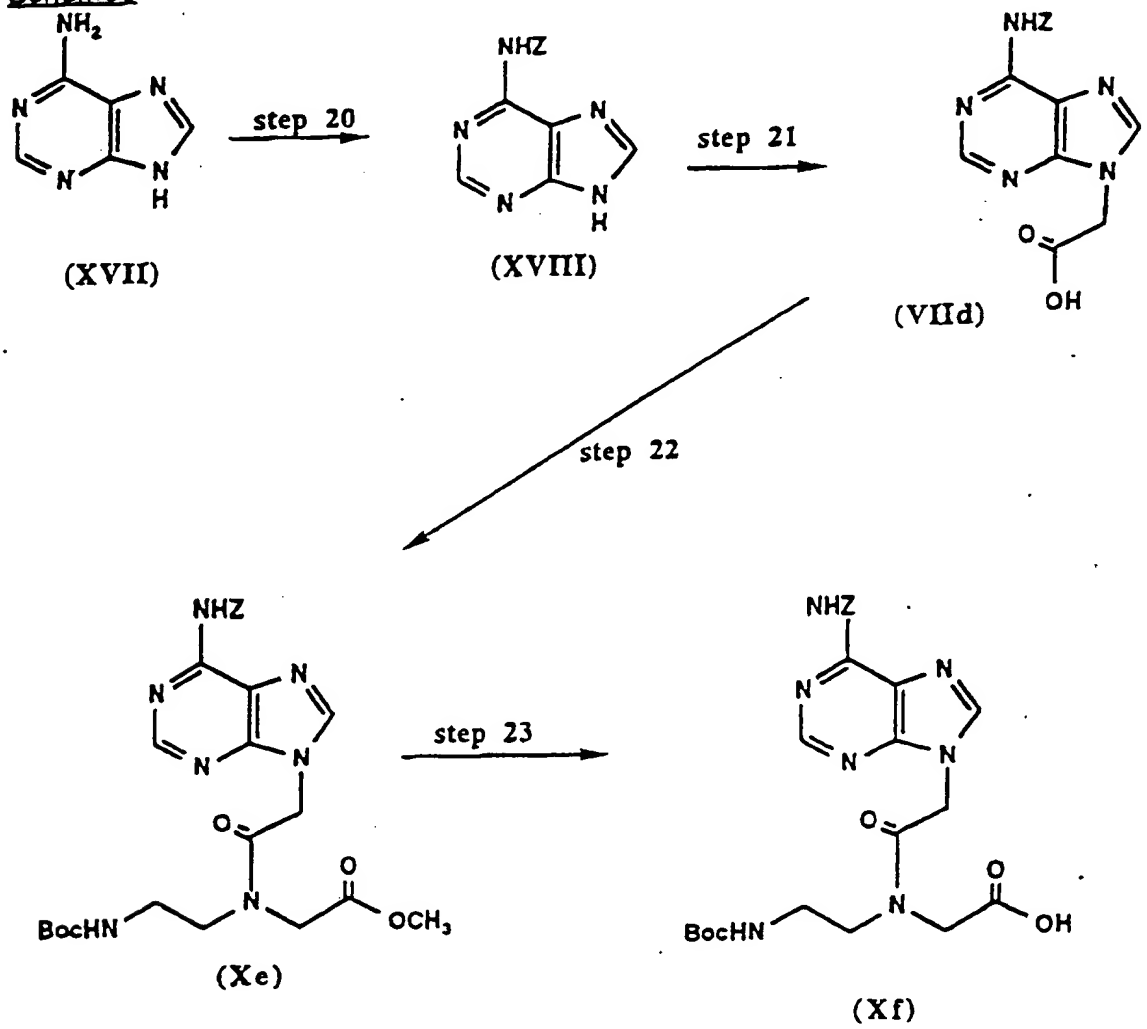
Scheme 5



Scheme 5 depicts a synthesis of the monomer of formula (X) in which R^1 and R^3 are hydrogen, R^2 is Boc, R^4 is methyl, and B is 6-O-benzyl-2-N-(benzyloxycarbonyl)-guanine (formula (Xd)). In step 16 the commercially available 2-amino-6-chloropurine (formula (XV)) is converted to the compound of (formula (XVI)) as described by M. MacCoss et al. in Tetrahedron Lett. 1985, 26, 1815. In step 17 the compound of formula (XVI) is alkylated with allyl bromide at the 9 position to give the compound of formula (IX) where B is protected guanine, formula (IXa). In step 18 the alkene moiety of the compound of (formula (IXa)) is oxidatively cleaved by treatment with sodium periodate in the presence ruthenium tetraoxide at ca. 25°C as described by Carlsen et al in J. Org. Chem. 1981, 46, 3936, to give the carboxylic acid which is methylated with diazomethane to give the particular compound of formula (VIIc). In step 19 the compound of formula (VII), i.e formula

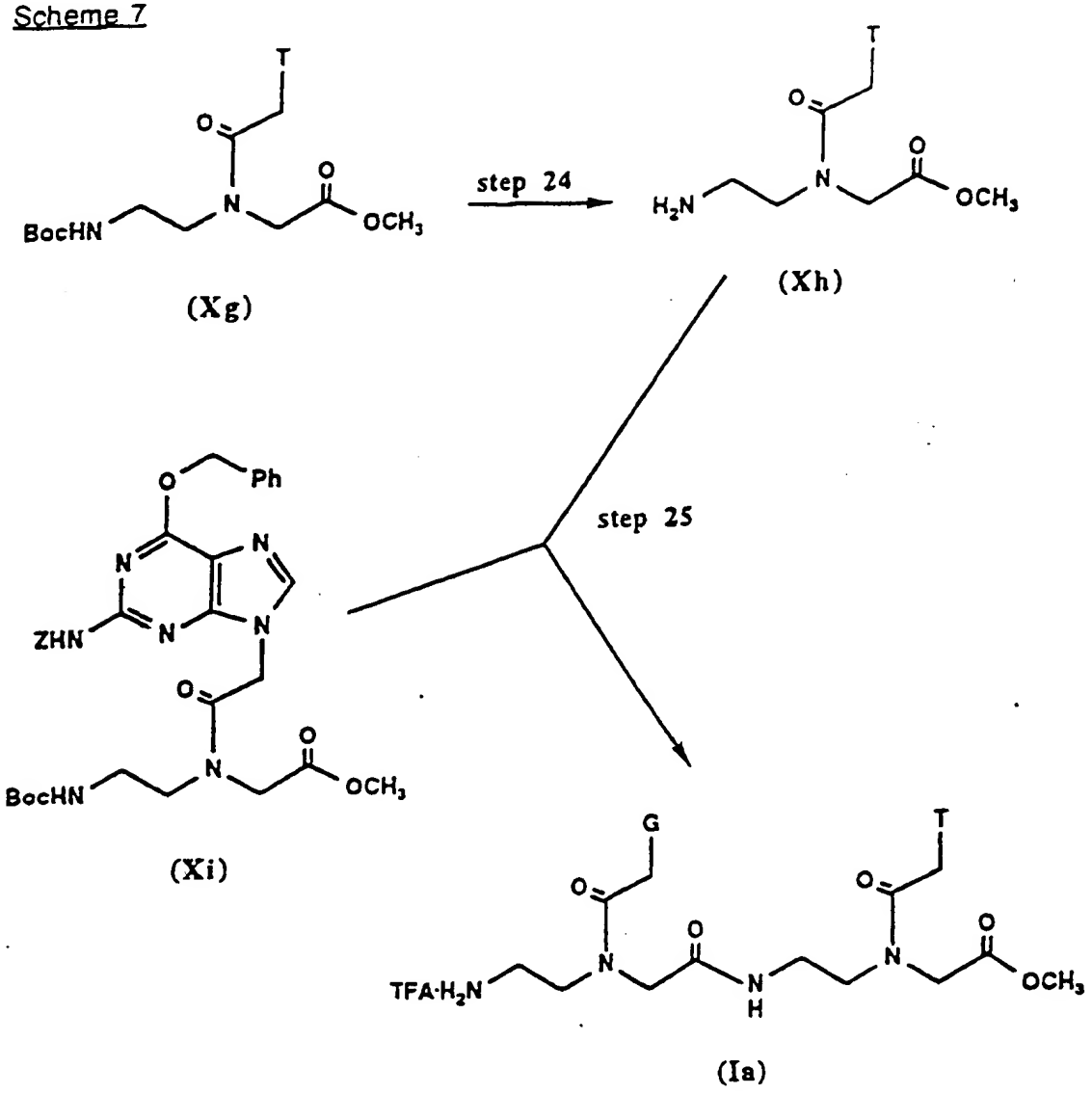
(VII), i.e. formula (VIIc) is first hydrolysed to the carboxylic acid, then activated with BOP in dimethylformamide and reacted with the compound of formula (Va) to give the compound of formula (Xd). The compound of formula (Xd) is referred to as the Bn-Z protected Geg monomer methyl ester.

Scheme6



Scheme 6 depicts a synthesis of the monomer (X) in which R^1, R^3 and R^4 are hydrogen, R^2 is Boc and B is 6-N-benzyloxycarbonyladenine (formula (Xf)). In step 20 the exocyclic amine group of adenine (A-H, formula (XVII)) is protected with the benzyloxycarbonyl group (Z) to give the compound of formula (XVIII). In step 21 the compound of formula (XVIII) is reacted with tert-butyl bromoacetate, which is followed by removal of the t-butyl group with strong acid (trifluoroacetic acid) to give (VIId). In step 22 the compound of formula (VIId) is activated with BOP in dimethylformamide, and reacted with the compound of formula (Va) to give the compound of formula (Xe). In step 23 the methyl ester of the compound of formula (Xe) is hydrolysed with aqueous sodium hydroxide to give the compound of formula (Xf), which is referred to as the Z protected Aeg monomer.

Scheme 7



Scheme 7 is a more detailed description of step 7 of scheme 1, and depicts a method for making the compound of formula (I) in which n is 1, and reading left to right Q is hydrogen, R^1 is hydrogen, B is guanine, R^3 is hydrogen, R^1 is hydrogen, B is thymine, R^3 is hydrogen, and J is methoxy (formula (1a)). In step 24 the compound of formula (Xg) is treated with hydrogen chloride in dioxane to give the compound of formula (Xh) as the hydrochloride salt. In step 25 the carboxyl group of the compound of formula (Xi), the Bn-Z protected Geg monomer, is activated with HBTU and reacted with the compound of formula (Xh). The protecting groups are removed by first treating with triflouroacetic acid, followed by hydrogen fluoride to give the compound of formula (Ia). The compound of formula (Ia) is referred to as the Geg-Teg methyl ester.

The compounds of formula (I) may also be prepared by the solid phase method as described by Merrifield et al. in J. Am. Chem. Soc. 1963, 85, 2149, and Science 1986, 232, 341.

Figure 1 outlines a synthesis of the compound of formula (I) in which n is 5, Q is hydrogen, all R^1 an R^3 are hydrogen and all B are thymine, and J is lysine (C-terminal amide). In step a of figure 1 the Teg monomer is coupled to the free alpha-amino group of lysine which is bound to a MBHA resin. After coupling is complete the resin is washed. In step b the Boc group is removed by treatment of the resin with trifluoroacetic acid in methylene chloride. After the de-Boc reaction is complete the resin is washed and a second coupling can take place. After a total of six coupling and deprotection cycles, the resin is dried under vacuum, and in step c the resin is treated with hydrogen fluoride which cleaves the product from the resin to give the compound of formula (Ib).

Figure 2 depicts an assay to show effective binding of a test compound nucleoside base oligomer of the invention of formula (I) to genetic material employing the enzyme RNase H. In this assay 3H labeled poly rA (RNA strand) is allowed to bind to its complementary DNA strand, dT (25 to 30 bases in length). After ca. 30 minutes the compound of formula (I), in particular where n=at least 5, is added at various concentrations, and the mixture is incubated for ca. 30 minutes, at which time the compound of formula (I), in particular where n= at least 5, binds to the poly rA strand by displacing the poly dT strand. The enzyme RNase H (from Hela cells) is then added to the mixture. RNase H will cleave the RNA strand of a RNA-DNA duplex, but not the RNA strand of a RNA-(Ia) duplex. Therefore only the portion of the poly rA strand which is bound to the dT strands will be cleaved into smaller fragments, and the portion of poly rA which is bound to the nucleoside base oligomer (formula (Ia)) will remain in tact. After ca. 30 minutes t-RNA and acid is added which precipitates the larger pieces of the poly rA, and the radioactivity remaining in the supernatant is counted. A decrease in radioactivity in the supernatant is a measure of the binding of the nucleobase oligomer of the invention over dT.

Figure 3 shows the results for the assay of Figure 2 for the nucleoside base oligomer of formula (Ia). In this graph the Y axis is the radioactivity in the supernatant and the X axis is the concentration of the compound of formula (Ia). As can be seen from the graph increasing the concentration of the compound of formula (Ia) results in a strong decrease in the radioactivity in the supernatant. At a

concentration of 2 μM the compound of formula (Ia) has nearly totally displaced the dT strand from the poly rA strand. The addition of the compound of formula (Ia) had no effect on the RNase H cleavage of the RNA strand of an RNA-DNA duplex in which the RNA strand was not complementary to the base sequence of the compound of formula (Ia). Therefore the binding of the compound of formula (Ia) to RNA is sequence dependent.

Another assay can be used to measure the binding of the test compound nucleoside base oligomer of the invention of formula (I) to double stranded (ds) DNA by measuring inhibition of a restriction enzyme which cleaves the ds DNA near or within the binding site of the test comound. In this assay the test compound of formula (I) is allowed to bind to the target ds DNA which contains a complementary sequence to the test compound. After the test compound has been given time to bind to the ds DNA a restriciton enzyme is added and the amount of cleavage is measured. A second restriction enzyme site removed from the test compound binding site is also within the ds DNA sequence and this site is used as an internal control. A decrease in ds DNA cleavage near the test compound binding site is a measure of increased binding of the test compound.

Pharmaceutical Compositions

It will further appreciated that the amount of a compound of the invention required for use in treatment will vary not only with the particular compound selected but also with the route of administration, the nature of the condition being treated and the degree and condition of the patient and will be ultimately at the discretion of the attendant physician or veterinarian. In general however a suitable dose will be in the range of from about 1 to 75 mg/kg of bodyweight per day, such as about 0.01 to about 50 mg per kiolgram body weight of the recipient per day, preferably in the range of 0.025 to 40 mg/kg/day.

The desired dose may be presented in a single dose or as divided doses administered at appropriate intervals, for example as two, three, four or more sub-doses per day.

The compound is conveniently administered in unit dosage form. For example containing 0.1 to 1000, conveniently 0.2 to 500, most conveniently 0.4 to 250 mg of active ingredient per unit dosage form.

Formulations of the present invention, for medical use, comprise an active compound, i.e., a compound of formula (I), together with an acceptable carrier therefof and optionally other therapeutically active ingredients. The carrier must be pharmaceutically acceptable in the sense of being compatible with the other ingredients of the formulation and not deleterious to the recipient thereof. The present invention, therefore, further provides a pharmaceutical formulation comprising a compound of formula (I) together with a pharmaceutically acceptable carrier thereof. The formulations include those suitable for oral, rectal or parenteral (including subcutaneous, intramuscular and intravenous) administration. Preferred are those suitable for oral or parenteral administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of pharmacy. All methods include the step of bringing the active compound into association with a carrier which constitutes one or more accessory ingredients. In general, the formulations are prepared by uniformly and intimately bringing the active compound into association with a liquid carrier or a finely divided solid carrier and then, if necessary, shaping the product into desired unit dosage form.

Formulations of the present invention suitable for oral administration may be presented as discrete units such as capsules, cachets, tablets or lozenges, each containing a predetermined amount of the active compound; as a powder or granules; or a suspension or solution in an aqueous liquid or non-aqueous liquid, e.g., a syrup, an elixir, an emulsion or a draught.

A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets may be prepared by compressing in a suitable machine the active compound in a free-flowing form, e.g., a powder or granules, optionally mixed with accessory ingredients, e.g., binders, lubricants, inert diluents, surface active or dispersing agents. Molded tablets may be made by molding in a suitable machine, a mixture of the powdered active compound with any suitable carrier.

A syrup or suspension may be made by adding the active compound to a concentrated, aqueous solution of a sugar, e.g., sucrose, to which may also be added any accessory ingredients. Such accessory ingredient(s) may include flavoring, an agent to retard crystallization of the sugar or an agent to increase the solubility of any other ingredient, e.g., as a polyhydric alcohol, for example, glycerol or sorbitol.

Formulations for rectal or vaginal administration may be presented as a suppository with a conventional carrier, e.g., cocoa butter or Witepsol 155 (trademark of Dynamite Nobel Chemical, Germany, for a suppository base).

Formulations suitable for parenteral administration conveniently comprise a sterile aqueous preparation of the active compound which is preferably isotonic with the blood of the recipient. Such formulations suitably comprise a solution or suspension of a pharmaceutically and pharmacologically acceptable acid addition salt of a compound of the formula (I) that is isotonic with the blood of the recipient. Thus, such formulations may conveniently contain distilled water, 5% dextrose in distilled water or saline and a pharmaceutically and pharmacologically acceptable acid addition salt of a compound of the formula (I) that has an appropriate solubility in these solvents, for example the hydrochloride. Useful formulations also comprise concentrated solutions or solids containing the compound of formula (I) which upon dilution with an appropriate solvent give a solution suitable for parental administration above.

In addition to the aforementioned ingredients, the formulations of this invention may further include one or more optional accessory ingredient(s) utilized in the art of pharmaceutical formulations, e.g., diluents, buffers, flavoring agents, binders, surface active agents, thickeners, lubricants, suspending agents, preservatives (including antioxidants) and the like.

EXAMPLES

Example 1

Methyl N-(2-tert-butyloxycarbonylaminoethyl)glycinate
(Formula (V) $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl)

To a solution of N-tert-butyloxycarbonylglycinal (15 g, 97 mmol, freshly prepared according to Thompson et al) in 400 mL of methanol (anhydrous) under nitrogen atmosphere is added glycine methyl ester hydrochloride (15.4 g, 122 mmol), sodium acetate (16.9g, 205.6 mmol), and 80 g of powdered 4 A molecular sieves. After stirring for ca. 5 minutes sodium cyanoborohydride (12.9 g, 205.6 mmol) is added in about 4 portions, and the reaction mixture is stirred for ca 6 hours. The mixture is filtered and the filtrate concentrated. The residue is partitioned between Chloroform (250 mL) and half saturated aqueous sodium bicarbonate (250 mL). The aqueous phase is extracted with chloroform, and the combined organics are washed with brine, dried over sodium sulfate, and concentrated. The residue is purified by high vacuum kugelrohr distillation to give the title compound as an oil; 9.37 g, 41% yield. 1 H NMR (300MHz, DCl_3): 4.97 (br s, 1 H), 3.71 (s, 3 H), 3.38 (s, 2 H), 3.18 (q, J = 6.5, 2 H), 2.72 (t, J = 6.5, 2 H), 1.42 (s, 9 H). mass spectrum: m/e calculated (M+H) = 233, observed = 233.

Example 2

(S)-Methyl N-(2-tert-butoxycarbonylaminoethyl)-2-(amino)propionate
(Formula (V): $R^1 = H$, R^2 = Boc, R^3 = (S)-methyl, R^4 = methyl

To a solution of N-tert-butoxycarbonylglycinal (6.33 g, 39.7 mmol, freshly prepared according to Thompson et al.) in 160 mL of methanol (anhydrous) under a nitrogen atmosphere is added (L)-alanine methyl ester hydrochloride (5.55 g, 39.7 mmol), sodium acetate (6.51 g, 79.4 mmol) and 40 g of freshly activated powdered 4 A molecular sieves. After stirring for ca. 2 min, sodium cyanoborohydride (5.00 g, 79.5 mmol) is added in one portion. The reaction mixture is stirred at room temperature for 1.5 h then filtered and concentrated to a solid. The solid is partitioned between ethyl acetate (500 mL) and half saturated aqueous sodium bicarbonate (200 mL). The organics are dried

over sodium sulfate, filtered then concentrated to an oil. The oil is purified by flash chromatography to afford the title compound (3.60 g, 37% yield). 1 H NMR (300MHz, DCl_3): 5.04(br s, 1H), 3.73 (s, 3H), 3.35 (q, J = 7 Hz, 1H), 3.30-3.11 (m, 2H), 2.80-2.72 (m, 1H), 2.63-2.55 (m, 1H), 1.45 (s, 9H), 1.30 (d, J = 7Hz, 3H); mass spectrum: m/e calculated (M+H) = 247, observed = 247.

Example 3

N-(2-tert-butyloxycarbonylaminoethyl)-N-(1-thyminylacetyl)aminoacetic acid (Formula (X): $R^1 = R^3 = H$, R^2 = Boc, $R^4 = H$, B = thymine)

(a) Methyl N-(2-tert-butyloxycarbonylaminoethyl)-N-(1-thyminylacetyl)aminoacetate (Formula (X) $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, B = thymine).

To a solution of the compound of formula (V) ($R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, 3.00g,. 12.91 mmol) in 90 mL of DMF (anhydrous) is added 1-carboxymethylthymine (formula(VII)), R5 = H, B = T, 2.97 g, 16.14 mmol, see A. S. Jones et al Tetrahedron 1973, 29, 2293), BOP (7.13 g, 16.14 mmol), HOBt (2.18 g, 16.14 mmol), and triethylamine (3.00 mL, 25.83 mmol). After stirring for ca. 4 h the reaction mixture is diluted with 200 mL of half saturated brine and extracted with ethyl acetate. The combined organics are washed with 1 N aqueous hydrochloric acid, saturated aqueous sodium bicarbonate, brine, dried over magnesium sulfate, and concentrated. The resulting residue is chromatographed on silica gel (9:1 ethyl acetate:hexane) to give the title compound as a white solid: 3.65 g, 71 % yield. 1H NMR (300 MHz, CDCl3): 7.02 (s, 0.25 H), 6.95 (s, 0.75 H), 5.51 (br s, 1 H), 4.5 (s, 1.5 H), 4.42 (s, 0.5 H), 4.20 (s, 0.5 H), 4.05 (s, 1.5 H), 3.81 (s, 0.75 H), 3.75 (s, 2.25 H), 3.52 (t, J = 5.7, 2 H), 3.39 (m, 2 H), 1.91 (s, 3 H), 1.44 (s, 9H); mass spectrum: m/e calculated (M+H) = 399, observed = 399.

(b) N-(2-tert-butyloxycarbonylaminoethyl)-N-(1-thyminylacetyl)aminoacetic acid (Formula (X): $R^1 = R^3 = H$, $R^2 =$ Boc, $R^4 = H$, B = thymine).

A solution of the compound of formula (X) ($R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, B = thymine, 3.10 g, 7.71 mmol) in 70 mL of THF and 20 mL of water is cooled on ice. To this is added 1 N aqueous lithium hydroxide (15 mL, 15 mmol) and the reaction mixture is stirred for ca. 30 minutes. The pH of the mixture is the adjusted to ca. 2 with solid sodium bisufate. The solution is diluted with 15 mL of water, saturated with sodium chloride, and extracted with ethyl acetate. The combined organics are dried over magnesium sulfate and concentrated. The residue is dissolved in 50 mL of 1:1 acetonitrile:water and lyophilized to give the title compound as a white powder; 2.41 g, 81 % yield. 1H NMR (300 MHz, d4-methanol): 7.30 (s, 0.66 H), 7.26 (s, 0.33), 4.74 (s, 1.33 H), 4.56 (s, 0.66 H), 4.27 (s, 0.66 H), 4.10 (s, 1.33 H), 3.51 (m, 2H), 3.20 (m, 2H), 1.87 (s, 3H), 1.44 (s, 9H), spectrum m/e calculated (m + H) = 385, observed (m + H) = 385.

Example 4

Z-Ceg Monomer

(Formula (X), $R^1 = R^3 = R^4 = H$, R^2 = Boc, B = 4-N-benzyloxycarbonylcytosine)

(a) 4-N-Benzyloxycarbonylcytosine

(Formula (XIV) where Z=4-N-benzoyloxy-carbonyl).

To a suspension of cytosine (5 g, 45 mmol) in 90 ml of pyridine (anhydrous), under a nitrogen atmosphere, in an ice bath, is added benzylchloroformate (8 mL, 56 mmol) dropwise. The mixture is brought to room temperature and stirred for ca. 16 hours. To the mixture is then added 4-dimethylaminopyridine (2.75 g, 22.5 mmol) and more benzylchloroformate (8 mL, 56 mmol). After stirring a total of ca. 40 hours, the reaction mixture is poured into 200 mL of ice-water and stirred for 5 minutes. The resulting white solid is filtered, washed with water, dichloromethane, and dried under vacuum, to give the title compound as a white solid: 7.66 g, 69% yield. 1H NMR (300 MHz, d6-DMSO): 11.1 (br s, 1 H), 7.80 (d, 1 H), 7.37 (m, 5H), 6.92 (d, 1H), 5.18 (s, 2H).

(b) 1-[(tert-Butoxycarbonyl)methyl]-4-N-benzyloxycarbonylcytosine (Formula (VII):R^5 = t-butyl, B = 4-N-benzyloxycarbonylcytosine).

To a suspension of 4-N-benzyloxycarbonylcytosine (3.39 g, 13.84 mmol) in 25 ml of DMF (anhydrous), under a nitrogen atmosphere, is added cesium carbonate (4.96 g, 15.22 mmol). The mixture is stirred for ca. 13 min. and then tert-butyl bromoacetate (2.46 mL, 15.22 mmol) is added dropwise, and the mixture is stirred for ca. 4 hours. The resulting solids are filtered and washed with ethyl acetate. The filtrate is partially concentrated and then partitioned between ethyl acetate and dilute brine. The organic phase is washed with dilute brine, water, saturated aqueous brine, dried over sodium sulfate and concentrated, to yield a yellow foam, which is purified by crystallization (dichloromethane/hexane) to give the title compound as white crystals: 2.75 g, 55% yield. 1NMR (300 MHz, CDCl3): 7.52 (s, 1 H), 7.50 (d, 1 H), 7.36 (s, 5 H), 7.22 (d, 1 H), 5.20 (s, 2 H), 4.50 (s, 2 H), 1.4(s, 9 H).

(c) 1-(Hydroxycarbonylmethyl)-4-N-benzyloxycarbonylcytosine (Formula (VII):R^5 = H, B = 4-N-benzyloxycarbonylcytosine).

To a solution of the compound of formula (VII) (6 g, 16.71 mmol, R^5 = t-butyl, B = 4-N-benzyloxycarbonylcytosine) in 45 ml of dichloromethane (anydrous), under a nitrogen atmosphere, is added anisole (20 mL, 184 mmol), followed by trifluoroacetic acid (50 mL, 649 mmol). The reaction mixture is stirred for ca. 4 hours, and then concentrated to dryness. To the residue is added toluene, and the solution is concentrated to dryness. This process is repeated two more times. The residue is dried under vacuum, and then triturated with dichloromethane. The resulting solids are filtered and washed with dichloromethane to give the title compound as a white solid; 6.38 g, 81% yield, as a 1:1 complex with trifluoroacetic acid. 1H NMR (300 MHz, d6-DMSO): 8.02 (d, J = 7.3, 1 H), 7.38 (m, 5 H), 7.01 (d, J = 7.3, 1 H), 5.18 (s, 2 H), 4.51 (s, 2 H); mass spectrum: m/e calculated (M + H) =304, observed (M + H) = 304.

(d) Z-C Monomer methyl ester
(Formula (X): $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, B = 4-N-benzyloxycarbonylcytosine).

To a solution of the compound of formula (V) (0.59 g, 2.54 mmol, $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl) in 8 ml of DMF (anhydrous), under a nitrogen atmosphere, is added BOP (0.94 g, 2.1 mmol), HOBt (0.29 g, 2.1 mmol) and a solution of the compound of formula (VII) (0.93 g, 2.23 mmol, R5 = H, B = 4-N-benzyloxycarbonylcytosine) and triethylamine (1.41 mL, 10.15 mmol) in 8 ml of DMF (anhydrous). After stirring for ca. 3 hours the reaction mixture is concentrated, and the residue partitioned between 100 mL of ethyl acetate, and 50mL of 0.5 N aqueous hydrochloric acid. The organic phase is washed with half saturated aqueous sodium bicarbonate and saturated aqueous brine. Crystallization from the ethyl acetate solution affords the title compound as white crystals:0.62 g, 59% yield. 1H NMR (300 MHz, CDCl3): 7.65 (br s, 1 H), 7.62 (d, 0.3 H), 7.58 (d, 0.7 H), 7.35 (s, 5 H), 7.22 (d, 1 H), 5.55 (t, 1 H), 5.19 (s, 2H), 4.70 (s, 1.4H), 4.55 (s, .6H), 4.30 (s, .6H), 4.05 (s, 1.4 H), 3.78 (s, 0.8 H), 3.70 (s, 2.2 H), 3.55 (t, 1.4 H), 3.50 (t, 0.6 H), 3.32 (q, 1.4 H), 3.22 (q, 0.6 H), 1.40 (s, 9 H); mass spectrum: m/e calculated (M + H) = 518, observed (M + H) = 518.

(e) Z-Ceg Monomer
(Formula (X): $R^1 = R^3 = R^4 = H$, R^2 = Boc, B = 4-N-benzyloxycarbonylcytosine).

To a suspension of the compound of formula (X) (2.36 g, 4.56 mmol, $R^1 = R^3$ = H, R^2 = Boc, R^4 = methyl, B = 4-N-benzyloxycarbonylcytosine) in 70 mL of THF-water (1/1), in an ice bath, is added 1 N aqueous sodium hydroxide (14 mL, 14 mmol). After stirring ca. 10 minutes, the reaction mixture is partitioned between 80 mL of ethyl acetate, and 90 mL of dilute brine. The aqueous phase is washed with ethyl acetate, acidified with saturated aqueous sodium bisulfate, saturated with sodium chloride, and extracted ethyl acetate. The organic phases are dried over magnesium sulfate and partially concentrated. Crystallization occurs and the heterogenous solution is diluted with chloroform, ethyl acetate, and filtered to give the title compound as white crystals: 2.02 g, 88% yield. 1H NMR (300 MHz, d6-DMSO): 10.78 (br s, 1 H), 7.89 (d, 0.65 H), 7.86 (d, 0.35 H), 7.36 (m, 5 H), 7.02 (d, 0.65 H), 6.99 (d, 0.35 H), 6.93 (t, (s, 0.7 H), 6.74 (t, .35H), 5.18 (s, 2H), 4.80 (s, 1.3H), 4.60 (s, .7H),

4.20 (s, 0.7 H), 3.97 (s, 1.3 H), 3.38 (br t, 1.3 H), 3.28 (br t, 0.7 H), 3.18 (br q, 1.3 H), 3.00 (br q, 0.7 H), 1.35 (s, 9 H); mass spectrum: m/e calculated (M + H) = 504, observed (+ H) = 504.

Example 5

Bn-Z-Geg monomer methyl ester
(Formula (X): $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine)

(a) 9-N-allyl-6-O-benzylguanine (Formula (IXa).

To a solution of 6-O-benzylguanine (0.930 g, 3.85 mmol) (MacCoss) in 10 mL of dry DMF at room temperature is added potassium carbonate (2.66 g, 19.3 mmol) and 18-crown-6 (1.02 g, 3.85 mmol). After 0.25 h, allyl bromide (0.367 mL, 4.24 mmol) is added in one portion. The resulting solution is vigorously stirred for 1 h. The mixture is partitioned between 125 mL of ethyl acetate and 50 mL of water. The organics are washed with brine, dried over sodium sulfate, filtered and concentrated. The residue is purified by radial chromatography on silica gel (hexane/ethyl acetate, gradient elution) to afford the title compound as an oil (0.582 g, 54%). 1H NMR (300 MHz, CDCl3): 7.60 (s, 1H), 7.52-7.48 (m, 2H), 7.38-7.7 (m, 3H), 6.05-5.92 (m, 1H), 5.56 (s, 2H), 5.26 (dd, J = 3.4, 0.3 Hz, 1H), 5.13 (dd, J = 5.7, 0.3 Hz, 1H), 5.05 (br s, 2H), 4.65 (ddd, J = 1.9, 0.3, 0.3 Hz, 2H); mass spectrum: m/e calculated (M+H) = 282, observed = 282.

(b) 9-N-allyl-6-O-benzyl-2-N-(benzyloxycarbonyl)guanine (Formula (IX): B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine).

To a solution of 9-N-allyl-6-O-benzylguanine (0.170 g, 0.604 mmol) in 5 mL of THF at room temperature is added 18-crown-6 (0.319 g, 1.21 mmol) and N-(benzyloxycarbonyl)imidazole (0.611 g, 3.02 mmol, Watkins). After 5 min, potassium hydride (35% in oil, 0.173 g, 1.51 mmol) is added dropwise. The resulting solution is maintained at room temperature for 1 h. The mixture is then partitioned between 100 mL of ethyl acetate and 50 mL of water. The organics are washed with brine, dried over sodium sulfate, filtered then concentrated. The residue is purified by radial chromatography on silica gel (hexane/ethyl acetate, gradient elution), to afford the title compound as an oil;

().236 g, 93%). 1H NMR (300 MHz, CDCl3): 7.82 (br s, 1H), 7.76 (s, 1H), 7.53-7.42 (m, 2H), 7.42-7.28 (m, 8H), 6.04-5.91 (m, 1H), 5.59 (s,2H), 5.28 (dd, J = 3.5, 0.4 Hz, 1H), 5.25 (s, 2H), 5.20 (dd, J = 5.6, 0.3 Hz, 1H), .73 (ddd, J=1.9, 0.4, 0.4 Hz, 2H).

(c) 6-O-benzyl-2-N-(benzyloxycarbonyl)-9-N-[(methoxycarbonyl)methyl]guanine (Formula (VII):R^5 = methyl, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine).

To a solution of the compound of formula (IX) (0.230 g, 0.554 mmol, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)- guanine) in 2 mL of carbon tetrachloride, 2 mL of acetonitrile and 3 mL of water is added sodium periodate (0.474 g, 2.21 mmol) followed by ruthenium(III) chloride hydrate (0.010 g, 0.048 mmol). After 5 h at room temperature an additional amount of ruthenium (III) chloride hydrate (0.010 g, 0.048 mmol) is added. The resulting solution is vigorously stirred at room temperature for 15 h. The mixture is partitioned between 100 mL of methylene chloride and 25 mL of water. The organics are dried over sodium sulfate, filtered then concentrated. The residue is partially purified by radial chromatography on silica gel (methanol/methylene chloride/acetic acid, gradient elution) to afford the free acid. The free acid is azeotropically dried with toluene and the residue dissolved in methanol. A solution of diazomethane in ethyl ether is added dropwise until the yellow color persists. The volatiles are removed under reduced pressure and the resulting oil is purified by radial chromatography on silica gel (hexane/ethyl acetate, gradient elution) to afford the title compound as an oil (0.061 g, 25%). 1H NMR (300 MHz, CDCl3): 7.85 (s, 1H), 7.54-7.51 (m, 2H), 7.45-7.30 (m, 8H), 5.61 (s, 2H), 5.27 (s, 2H), 4.97 (s, 2H), 3.79 (s, 3H).

(d) Bn-Z-Geg monomer methyl ester (Formula (X): R^1 = R^3 = H, R^2 = Boc, R^4 = methyl, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine).

To a solution of the compound of formula (VII) (0.055 g, 0.12 mmol, R^5 = methyl, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine) in 2 mL of THF and 1 mL of water at room temperature is added lithium hydroxide monohydrate (0.015 g, 0.37 mmol). After 0.5 h the reaction is acidified with 2 mL of 5% aqueous hydrochloric acid and extracted with ethyl acetate. The organics are dried over sodium sulfate, filtered then concentrated to a white

powder (0.050 g, 94%). To a solution of the resulting free acid and the compound of formula (V) (0.050 g, 0.21 mmol, $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl) in 2 mL of dry DMF is added BOP (0.076 g, 0.17 mmol) and HOBt (0.023 g, 0.17 mmol). After 5 min, triethylamine (0.048 mL, 0.35 mmol) is added in one portion. The resulting solution is stirred at room temperature for 2 h. The mixture is partitioned between 100 mL of ethyl acetate and 50 mL of brine. The organics are washed with brine, dried over sodium sulfate, filtered then concentrated to an oil. The oil was purified by radial chromatography on silica gel (methanol/methylene chloride, gradient elution), to afford the title compound as an oil; 0.85 g, quantitative yield). 1H NMR (250 MHz, CDCl3): 7.96 (s, .75H), 7.89 (s, .25H), 7.69 (s, 1H), 7.49-7.46 (m, 2H), 7.38-7.27 (m, 8H), 6.27 (dd, J = 5.1, 5.1 Hz, 1H), 5.55 (s, 2H), 5.22 (s, 2H), 5.03 (s, 1.5H), 4.86 (s, .5H), 4.37 (s, .5H), 4.09, (s, 1.5H), 3.78 (s, .75H), 3.70 (s, 2.25H), 3.63 (m, 1.5H), 3.51 (m, .5H), 3.39 (m, 1.5H), 3.23 (m, .5H), 1.20 (s, 9H); mass spectrum: m/e calculated (M+H) = 648, observed = 648.

Example 6

Z-Aeg Monomer

(Formula (X): $R^1 = R^3 = R^4 = H$, R^2 = Boc, B = 6-N-enzyloxycarbonyladenine)

(a) 9-[(tert-Butyloxycarbonyl)methyl]-adenine
(Formula (VII): R^5 = tert-butyl, B = adenine).

To a suspension of adenine (3 g, 22.2 mmol) in 60 mL of DMF (anhydrous), under a nitrogen atmosphere, is added cesium carbonate (7.96 g, 24.42 mmol), followed by the dropwise addition of tert-butyl bromoacetate (4.3 mL, 26.64 mmol). The reaction mixture is stirred at room temperature for ca. 14 hours, concentrated to half the volume and partitioned between 125 mL of chloroform and 100 mL of dilute brine. The aqueous phase is extracted with chloroform, and the combined organic phases are washed with 30 mL of saturated brine. Upon partial concentration, crystallization occurs. The white crystals were washed with ethyl acetate and dried to afford the title compound: 3.05 g, 55% yield. 1H NMR (300 MHz, d6-DMSO): 8.12 (s, 1 H), 8.09 (s, 1 H), 7.24 (br s, 2 H), 4.93 (s, 2 H), 1.41 (s, 9 H).

(b) 6-N-(Benzyloxycarbonyl)-9-[(tert-butyloxycarbonyl)methyl]-adenine (Formula (VII) R^5 = tert-butyl, B = 6-N-benzyloxycarbonyladenine).

To a solution of N-(benzyloxycarbonyl)imidazole (0.16 g, 0.8 mmol, see J. Am. Chem. Soc., 1982, 104, 5702-5708) in 1.5 mL of 1,2-dichloroethane (anhydrous), under a nitrogen atmosphere, in an ice bath, is added triethyloxonium tetrafluoroborate (0.82 mL, 1.0 M in methylene chloride, 0.82 mmol). After the addition, the ice bath is removed and the mixture is stirred at room temperature for ca. 2 hours. To the reaction mixture is then added the compound of formula (VII) (0.050 g, 0.20 mmol, R5 = tert-butyl, B = adenine) as a solid. The reaction mixture is heated at 82 C for ca. 5 hours, and allowed to stand at room temperature for ca. 2 days. The mixture is concentrated and chromatographed twice on silica gel (first;ethyl acetate-hexane gradient, second; ethyl acetate-chloroform gradient) to give the title compound as a white solid: 0.040g, 53% yield. 1H NMR (300 MHz, CDC 3): 8.73 (s, 1H), 8.46 (br s, 1H), 7.96 (s, 1 H), 7.37 (m, 5 H), 5.28 (s, 2 H), 4.84 (s, 2 H), 1.47 (s, 9 H); mass spectrum: m/e calculated (M + H) = 384, observed (m + H) = 384.

(c) 6-N-(Benzyloxycarbonyl)-9-[(hydroxycarbonyl)methyl]-adenine (Formula (VII):R^5 = H, B = 6-N-benzyloxycarbonyladenine).

To a solution of the compound of formula (VII) (0.115 g, 0.30 mmol, R5 = tert-butyl, B = 6-N-benzyloxycarbonyladenine) in 5 mL of dichloromethane (anhydrous), under a nitrogen atmosphere, is added anisole (2.5 mL, 23 mmol), followed by trifluoroacetic acid (7 mL, 91.0 mmol). The reaction mixture is stirred for ca. 4 hr, and concentrated to dryness. The residue is azeotroped five times from chloroform, methanol and ethyl ether mixtures and dried under high vacuum overnight, to give the title compound as a 1:1 complex with trifluoroacetic acid: 0.134 g, 90% yield. 1H NMR (300 MHz, d4-3 H), 8.67 (s, 1H), 8.52 (s, 1H), 7.45 (dd, 2H), 7.38 (m, 3H), 5.35 (s, 2 H), 5.18 (s, 2 H); mass spectrum: m/e calculated (M + H) = 328, observed (m + H) = 328.

(d) Z-Aeg Monomer methyl ester
(Formula (X): $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl, B = 6-N-benzyloxycarbonyladenine).

To a solution of BOP (0.124 g, 0.28 mmol), HOBt (0.038g, 0.28mmol) and the compound of formula (VII) (0.130 g, 0.28 mmol, R5 = H, B = 6-N-benzyloxycarbonyladenine) in 1.5 mL of DMF (anhydrous), under a nitrogen atmosphere, is added triethylamine (0.188 mL, 1.35 mmol). After stirring the mixture for ca. 1.5 min, it is transferred via syringe into a solution of the compound of formula (V) (0.07 g, 0.27 mmol, $R^1 = R^3 = H$, R^2 = Boc, R^4 = methyl) in 1 mL of DMF (anhydrous), and the mixture is stirred for ca. 2 hours. To the reaction mixture is added BOP (0.124 g, 0.28 mmol) and the mixture is stirred for ca. 3 hours. The mixture is concentrated to dryness and partitioned between 25 mL of ethyl acetate, and 10 mL of 0.5 N hydrochloric acid containing 3 mL of saturated aqueous brine. The organic phase is washed with aqueous 0.5 N hydrochloric acid-brine, dilute aqueous sodium bicarbonate-brine, and saturated aqueous brine. The organics are concentrated and chromatographed on silica gel (first; ethyl acetate-hexane, second; methanol-ethyl acetate) to give the title compound as a white foam: 0.078 g, 53% yield. 1H NMR (300 MHz, CDCl3): 8.72 (s, 3 H), 8.14 (br s, 1H), 8.02 (s, 1H), 7.42 dd, 2H), 7.36 (m< 3H), 5.60 (br t, 1 H), 5.28 (s, 2 H), 5.14 (s, 1.6 H), 4.97 (s, 0.4 H), 4.29 (s, 0.4 H), 4.05 (s, 1.6 H), 3.82 (s, 0.75 H), 3.73 (s, 2.25 H), 3.64 (t, 1.6 H), 3.54 (t, 0.4 H), 3.38 (q, 1.6 H), 3.26 (q, 0.4H), 1.4 (s, 9H); mass spectrum: m/e calculated (m +H) = 541, observed (m + H) = 541.

(e) Z-Aeg Monomer
(Formula (IX): $R^1 = R^3 = R^4 = H$, R^2 = Boc, B = 6-N-benzyloxycarbonyladenine).

To a solution of the compound of formula (X) (0.070 g, 0.129 mmol, $R^1 = R^3 =$ H, R^2 = Boc, R^4 = methyl, B = 6-N-benzyloxycarbonyladenine) in 3 mL of 1:1 THF-water, in an ice bath, is added 1 N aqueous sodium hydroxide (0.388 mL, 0.388 mmol). After stirring for ca. 1 hour the reaction mixture is partitioned between 5 mL of ethyl acetate and 10 mL of water. The aqueous phase is washed with ethyl acetate, acidified with 3 mL of saturated aqueous sodium bisulfate, saturated with sodium chloride and extracted with ethyl acetate. The combined organic phases are dried over sodium sulfate and

concentrated to give the title compound as a white solid: 0.065 g, 95% yield. 1H NMR (300 MHz, d4 methanol): 8.58, (s, 0.6H), 8.57 (s, 0.4H), 8.27 (s, 0.6 H), 8.26 (s, 0.4 H), 7.46 (d, 2 H), 7.36 (m, 3 H), 5.39 (s, 1.4 H), 5.30 (s, 2 H), 5.21 (s, 0.6 H), 4.40 (s, 0.6 H), 4.12 (s, 1.4 H), 3.64 (t, 1.4 H), 3.47 (t, 0.6 H), 3.38 (t, 1.4H), 3.20 (t, 0.6H), 1.41 (s, 9H); mass spectrum: m/e calculated (m + H) = 527, observed (m + H) = 527.

Example 7

(S)-N-(2-tert-butoxycarbonylaminoethyl)-N-(1-thyminylacetyl)-2-(amino)propionic acid (Formula (X) R^1 = H, R^2 = Boc, R^3 = (S)-methyl, R^4 = H, B = thymine)

(a) (S)-Methyl N-(2-tert-butoxycarbonylaminoethyl)-N-(1-thyminylacetyl)-2-(amino)propionate (Formula (X) R^1 = H, R^2 = Boc, R^3 = (S)-methyl, R^4 = methyl, B = thymine).

To a solution of the compound of formula (V) (R^1 = H, R^2 = Boc, R^3 = (S)-methyl, R^4 = methyl, 1.7 g, 6.9 mmol) in 20 mL of DMF (anhydrous) is added 1-carbomethoxythymine (formula (VII) R5 = H, B = thymine, 1.4 g, 7.6 mmol) and 1,3-dicyclohexylcarbodiimide (1.5 g, 7.6 mmol). After stirring at room temperature for ca. 20 h, the reaction mixture is filtered through celite. The filtrate is concentrated then partitioned between ethyl acetate (300 mL) and a 1:1 mixture of brine and saturated aqueous sodium bicarbonate (150 mL). The organics are dried over sodium sulfate, filtered then concentrated on ca. 10 g of silica gel. The silica gel is placed on the top of a flash column packed with silica and then the column is eluted with methanol/methylene chloride (gradient elution) to afford the title compound (2.6 g, 91%): 1H NMR (300 MHz, $CDCl_3$): 9.3 (br s, 0.25H), 8.97 (br s, .75H), 6.99 (s, .25H), 6.93 (s, .75H), 5.57 (br t, J = 5Hz, 1H), 4.54 (m, 2H), 4.34 (q, J = 7Hz, 1H), 3.78 (s, .25H), 3.73 (s, .75H), 3.67-3.22 (m, 4H), 1.90 (s, 3H), 1.60 (d, J = 7Hz, 3H), 1.43 (s, 9H); mass spectrum: m/e calculated (M+H) = 413, observed = 413.

(b) (S)-N-(2-tert-butoxycarbonylaminoethyl)-N-(1-thyminylacetyl)-2-(amino)propionic acid (Formula (X): R^1 = H, R^2 = Boc, R^3 = (S)-methyl, R^4 = H, B = thymine).

A solution of the compound of formula (X) (R^1 = H, R^2 = Boc, R^3 = (S)-methyl, R^4 = methyl, B = thymine, 2.50 g, 6.06 mmol) om 30 mL of methanol and 15 mL of water is treated with sodium hydroxide water (0.29 g, 7.3 mmol). The reaction mixture is left at room temperature. After ca. 60 h, additional sodium hydroxide (0.12 g, 3.0 mmol) is added. After ca. 24 h, the methanol is removed under reduced pressure. The residual aqueous mixture is diluted with ethyl acetate (300 mL) and brine (100 mL). The pH of the aqueous phase is adjusted to ca. 2 with solid sodium bisulfate and the layers separated. The aqueous layer is back extracted with ethyl acetate (150 mL). The combined organics are dried over sodium sulfate, filtered then concentrated to a foam. The foam is dissolved in methylene chloride then added dropwise to vigorously stirred hexane (300 mL). The resulting precipitate is filtered and dried to afford the title compound as a white powder (2.12 g, 88% yield). 1H NMR (300 MHz, d6-DMSO): 11.29 (s, 1H), 7.35 (s, 1H), 6.93 (br s, .67H), 6.86 (br s, .33H), 4.68-4.56 (m, 2H), 4.31 (q, J = 7Hz, 7Hz, 3H), 3.2-2.93 (m, 4H), 1.73 (s, 1H), 1.39 (s, 9H), 1.34 (d, J=7Hz, 3H); mass spectrum: m/e calculated (M+H) = 399, observed = 399.

Example 8

Geg-Teg dimer
(Formula (I): $R^1$1 = $R^3$1 = $R^1$2 = $R^3$2 = H, B1 = T, B2 = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine, J = methoxy, Q = Boc)

(a) Teg monomer amine hydrochloride
(Formula (X): R^1 = R^3 = R^2 = H, R^4 = methyl, B = T).

The compound of formula (X) (0.093 g, 0.237 mmol, R^1 = R^3 = H, R^2 = Boc, B = T) is dissolved in 6 mL of 4 N hydrochloric acid in dioxane (Pierce). After 0.25 h, the reaction is concentrated to afford a white powder of the title compound (0.081 g, quantitative yield). 1H NMR (300 MHz, CDCl3): 7.33 (s, .33H), 7.26(s, .67H), 4.71 (s, .33H), 4.54 (s, .67H), 4.34 (s, .67H), 4.12 (s, .33H), 3.82-3.49 (m, 2H), 3.76 (s, 3H), 3.29-3.03 (m, 2H), 1.80 (s, 3H); mass spectrum: m/e calculated (M+H) = 299, observed = 299.

(b) Geg-Teg dimer
(Formula (I)in which n is 1, and reading left to right Q is Boc, R^1 is hydrogen, B is 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine, R^3 is hydrogen, R^1 is hydrogen, B is thymine, R^3 is hydrogen, and J is methoxy.

To a room temperature solution of the compound of formula (IX) (0.081 g, 0.12 mmol, R^1 = R^3 = H, R^2 = Boc, R^4 = methyl, B = 6-O-benzyl-2-N-(benzyloxycarbonyl)guanine) in 5 mL of THF and 2 mL of water is added lithium hydroxide monohydrate (0.008 g, 0.2 mmol). After 1 h, the reaction is quenched with solid sodium bisulfate then partitioned between 75 mL of ethyl acetate and 25 mL of brine. The aqueous phase is back extracted with 50 mL of ethyl acetate. The combined organics are dried over sodium sulfate, filtered then concentrated to a yellow powder (0.083 g, 0.13 mmol). To a solution of the resulting free acid (0.083 g, 0.13 mmol) and the compound of formula (V) (0.040 g, 0.12 mmol, R^1 = R^3 = R^2 = H, R^4 = methyl, B = T) in 1 mL of dry DMF is added a solution of HOBT/HBT (.45 M, 0.18 mmol). After 1 min, DIEA (0.063 mL, 0.36 mmol) is added in one portion. After 1 h at room temperature, the mixture is partitioned between 100 mL of ethyl acetate and 50 mL of brine. The aqueous layer is back extracted with an additional 50

mL of ethyl acetate. The combined organics are dried over sodium sulfate, filtered and concentrated. Radial chromatography on silica gel (methanol/methylenechloride, gradient elution) affords the title compound (0.065 g, 60%); mass spectrum: m/e calculated (M+H) = 914, observed = 914.

(c) Ceg-Teg dimer

(Formula (I) in which n is 1, and reading left to right Q is hydrogen, R^1 is hydrogen, B is guanine, R^3 is hydrogen, R^1 is hydrogen, B is thymine, R^3 is hydrogen, and J is methoxy.)

To a solution of trifluoroacetic acid (0.5 mL) and methylene chloride (1.0 mL) under a nitrogen atmosphere is added the fully protected guanine-thymine dimer (.044 g, .048 mmol). After 0.5 h at room temperature, the volatiles are removed in vacuum. After ca. 16 h under high vacuum, the resulting powder is dissolved in anisole (.020 mL) and hydrogen fluoride (ca. 5 mL) is added. After 1 h in an ice bath, the volatiles are removed under reduced pressure and the residue is dissolved in trifluoroacetic acid (ca. 5 mL) then concentrated. The residue is dissolved in water (5 mL) and acetonitrile (5 mL) then lyophilized to afford the guanine-thymine dimer methyl ester as a white powder (.028 g, quantitative); mass spectrum: m/e calculated (M+H) = 590, observed = 590.

Example 9

Ceg-Teg dimer

(Formula (I)) in which n is 1, and reading left to right Q is Boc, R^1 is hydrogen, B is 4-N-benzyloxycarbonylcytosine, R^3 is hydrogen, R^1 is hydrogen, B is thymine, R^3 is hydrogen and J is methoxy.

To a solution of Ceg monomer (0.030 g, 0.060 mmol) in 0.5 mL of anhydrous DMF is added HBTU and HOBt (0.133 mL of a 0.45 M solution in DMF, 0.60 mmol each), followed by DIEA (0.017 mL, 0.10 mmol). After stirring for 15 minutes the mixture is transferred, via syringe, to a solution of the compound of formula (V) (0.040 g, 0.12 mmol, $R^1 = R^3 = R^2 = H$, R^4 = methyl, B = T, 0.016 g, 0.050 mmol) and DIEA (0.017 ml, 0.100 mmol) in 1 ml of anhydrous DMF that was premixed for 25 minutes. The reaction is followed by HPLC and judged to be complete afer 60 minutes. The mixture is concentrated and

the residue dissolved in 15 mL of chloroform and washed with 0.3N hydrochloric acid, dilute sodium bicarbonate and saturated brine. The organic phase is dried over sodium sulfate, filtered and concentrated. Preparative thin layer chromatography on silica gel (3 plates, 20X20 cm, 2000 micron thickness, 11% methanol-chloroform as eluent) affords the title compound (0.019 g, 49%). 1H-NMR (300 MHz, CDCl3): Rotamers observed. Major rotamer: 7.6 (broad signal, 1H), 7.55 (broad signal, H), 7.35 (m, 5H), 7.2 (broad signal, 1H), 7.15 (s, 1H), 5.70 (broad signal, 1H), 5.15 (s, 2H), 4.7 (s, 2H), 4.6 (s, 2H), 4.05 (s, 4H), 3.75 (s, 3H), 3.7-3.15 (m, 8H), 1.75 (s, 3H), 1.35 (s, 9H).

Example 10

Aeg-Teg dimer

(Formula (I)) in which n is 1, and reading left to right 0 is Boc, R^1 is hydrogen, B is 6-N-benzyloxycarbonyladenine, R^3 is hydrogen, R^1 is hydrogen, B is thymine, R^3 is hydrogen and J is methoxy.

To a solution of Aeg monomer (0.012 g, 0.023 mmol) in 1 mL of anhydrous DMF is added HBTU and HOBt (0.050 mL of a 0.45 M solution in DMF, 0.023 mmol each), followed by DIEA (0.015 mL, 0.086 mmol). After stirring for 3 minutes, the mixture is transferred, via syringe, to a solution of the compound of formula (V) ($R^1 = R^3 = R^2 + H$, R^4 = methyl, B =T, 0.009 g, 0.022 mmol), and DIEA (0.015 ml, 0.086 mmol) in 0.8 mL of DMF that was premixed for 10 minutes. The reaction is followed by HPLC and judged to be complete after 30 minutes. The mixture is concentrated, the residue dissolved in chloroform and washed with dilute sodium bicarbonate and saturated brine. The organic phase is dried over sodium sulfate, filtered and concentrated. Preparative TLC chromotography on silica gel (2 plates, 20x20 cm, 2000 micron thickness, 15% methanol-chloroform as eluent) affords the title compound (0.008 g, 44%). 1H-NMR (300 MHz, MeOH-d4): Four rotamers observed. Major rotamer: 8.56 (s, 1H), 8.19 (s, 1H), 7.5-7.3 (m, 5H), 7.08 (d, J=1.2 Hz, 1H), 5.42 (s, 2H), 5.30 (s, 2H), 4.55 (s, 2H), 4.12 (s, 4H), 3.72 (s, 3H), 3.7-3.2 (m, 8H), 1.6 (s, 3H), 1.4 (s, 9H).

Example 11

(Teg) 6-1ys-NH_2

(Formula (I)) in which n is 5, 0 is hydrogen, all R^1 and R^3 are hydrogen, all B are thymine, and J is NH_2.

To MBHA resin 91.00 g, 0.25 meq/g, Peptides Interantional) prewashed with DMF is added t-butyloxycarbonyl-Ne-2-chlorobenzyloxycarbonyl-L-lysine (0.416 g, 1.00 mmol) in 2.22 mL of a HBTU/HOBt DMF solution 90.45 M in both HBTU and HOBt, referred to as the coupling solution) and DIEA (0.3 mL). The mixture is shaken gently for ca. 6 hours at room temperature. The reaction solution is removed by filtration, and the resin is washed with DMF. To the resin is added 2 mL of an acetic anhydride, DIEA, DMF solution (0.4/0.7/1.5 ratio, referred to as the capping solution). This is shaken gently for 0.5 hours. The resin is washed with DMF and methylene chloride, and is then treated wth 2 mL of a 1/1 trifluoroacetic acid/methylene chloride solution (referred to as teh de-Bocing solution) for 30 minutes. The reaction solution is removed by filtration and the resin is washed with a 15% solution of DIEA in methylene chloride, methylene chloride, and dried under vacuum, to give 1.101 g of dry C1Z-lys-MBHA resin.

To the Teg monomer (0.193 mg, 0.50 mmol) is added 1.1 mL of the coupling solution , 1 mL of DMF, and 0.15 mL of DIEA, and this mixture is allowed to stand for ca. 3 minutes. This solution is then added to the above resin (0.50 g, prewashed with DMF). The mixture is removed by filtration and the resin is washed with DMF, and treated with 2 mL of the capping solution for ca. 30 minutes. The reaction solution is removed by filtration and the resin is washed with DMF and methylene chloride. The resin is then treated with 2 mL of the de-Bocing solution for ca. 30 minutes and washed with 15% DIEA in methylene chloride, methylene chloride and DMF. This coupling-capping-de-Bocing cycle is repeated a total of six times. After the final de-Bocing step the resin is washed with methylene chloride and dried under vacuum to give 674 mg of resin. A portion of this resin (50 mg) is treated with hydrofluoric acid (ca. 5 mL) in the presence of anisole (ca. 0.5 mL) for ca. 50 minutes. The hydrofluoric acid is removed under vacuum, and the residue is taken up in trifluoroacetic acid and filtered. The trifluoroacetic acid solution is

concentrated and the residue purified by reverse phase HPLC (acetonitrile/water gradient) to give the title compound as a white solid (12 mg, 55% yield). mass spectrum: m/e (electrospray) calculated (M+H)=1743, observed = 1743.

Example 12

Rate of Strand Invasion into a poly rA·T_{25-30} heteroduplex by the compound of (Formula (I)) in which n is 5, Q is hydrogen, all R^1 and R^3 are hydrogen, all B are thymine and J is NH_2.

The ^3H poly rA·T_{25-30} heteroduplex was prepared as follows: 50 ml ^3H poly rA (5 µCi, 940 pmol nucleotide) and 100 pmol T_{25-30} (2500-3000 pmol nucleotide) were incubated in buffer A (40 mM Tris·HCl pH 7.5, 50 mM NaCl, 8 mN NaCl, 8 mM $MgCl_2$, 2 mM spermidine) in a 480 µl reaction at 70°C for 5 minutes followed by slow cooling to room temperature over ca. 1 hour and then placed at 15°C for 15 minutes. To the above solution 10 µl of the compound of formula (I) (in which n is 5, Q is hydrogen, all R^1 and R^3 are hydrogen, all B are thymine, and J is NH_2, 1000 pmol ODN, 6000 pmol nucleotide) is added and at various times (0, 5, 10, 15, 20, 30, 40, 50, 60 minutes) 40 µl are removed and 1 µl Hela cell nuclear extract (Bethesda Research Laboratories), as a source of human RNase H, is added. The reaction is incubated at 15C for 5 minutes and then terminated by the addition of 50 µl of 1µg/ml tRNA and 100 µl of 2 M HCl, 0.2 M sodium pyrophosphate. The solutions are placed on ice for ca. 10 minutes, and then centrifuged at 12,000 x g for 10 minutes at 4°C. The supernatant is removed and the amount of ^3H determined by scintillation counting. The extent of strand invasion is determined by comparing the ^3H in the supernatant for each time point to that of control reactions. A reaction which contained none of the compound of formula (I) was performed as above to determine the maximum amount of ^3H in the supernatant. A reaction, in which the PNA and T_{25-30} were added simultaneously to the reaction, was performed to determine the minimum amount to ^3H released (an additional way to determine the minumum ^3H released was to conduct the reaction in the absence of both T_{25-30} and the PNA; both approaches gave essentially identical amounts of ^3H in the supernatant.

What is Claimed is:

1. A nucleoside base oligomer comprising at least one purine nucleoside base bound to a backbone having at least one peptide bond.

2. The nucleoside base oligomer of Claim 1, wherein said at least one purine nucleoside base is adenine or an equivalent thereof.

3. The nucleoside base oligomer of Claim 1, wherein said at least one purine nucleoside base is guanine or an equivalent thereof.

4. The nucleoside base oligomer of Claim 1, wherein said oligomer further comprises at least one pyrimidine nucleoside base.

5. The nucleoside base oligomer of Claim 4, wherein said at least one pyrimidine base is thymine or an equivalent thereof.

6. The nucleoside base oligomer of Claim 4, wherein said at least one pyrimidine nucleoside base is cytosine or an equivalent thereof.

7. The nucleoside base oligomer of Claim 1, wherein said oligomer comprises at least 5 nucleoside bases or equivalents thereof.

8. The nucleoside base oligomer of Claim 7, wherein said oligomer comprises at least 3 different nucleoside bases selected from the group consisting of adenine, guanine, thymine or cytosine or equivalents thereof.

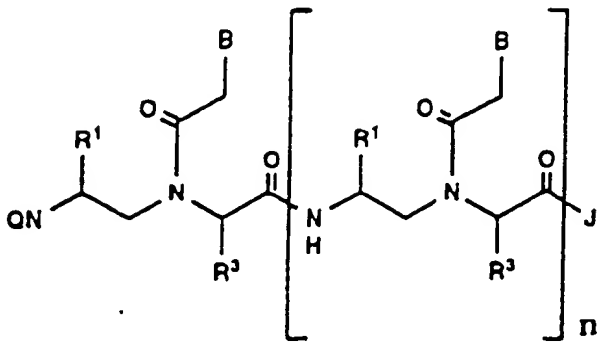
9. A method of affecting genetic material which comprises administering to the genetic material, the nucleoside base oligomer of Claim 1.

10. The method of Claim 9, wherein said method is a method of treating disease.

11. The method of Claim 9 wherein said method is a method of diagnosing a disease or condition.

12. The method of Claim 9 wherein said method is a method of recognizing said materials.

13. A nucleoside base oligomer of the following formula (I):



(I)

wherein

Q is an N-terminal blocking group;

J is a C-terminal blocking group or Q and J may together be a single bond;

n is at least 1;

R^1 is independently is hydrogen, benzyl, $-CH_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl, $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO(C_{1-4}$ alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkenyl, C_{2-8} alkynyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl) substituted amino;

R^3 is independently is hydrogen, benzyl, $-CH_2$-p-C_6H_4OH, $-CH_2$-indol-3-yl, $-CH_2CH_2CH_2CH_2NH_2$, $-CH_2CH_2CH_2NHC(NH)NH_2$, $-CH_2$-imidazol-4-yl, $-CH_2COOH$, $-CH_2COO$ (C_{1-4} alkyl), $-CH_2CH_2COOH$, $-CH_2CH_2COO(C_{1-4}$ alkyl), $-CH_2CONH_2$, $-CH_2CH_2CONH_2$, $-CH_2SH$, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkenyl, C_{2-8} alkynyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl, or di-(C_{1-4} alkyl) substituted amino;

B is independently a purine or pyrimidine nucleoside base providing that at lease one B is a purine nucleoside base,

or an acid or base addition salt thereof.

14. The nucleoside base oligomer of Claim 13 wherein at least one B is adenine.

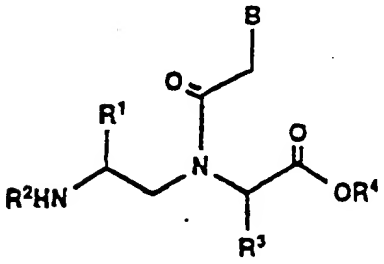
15. The nucleoside base oligomer of Claim 13 wherein at least one B is guanine.

16. The nucleoside base oligomer of Claim 13 wherein at least one B is thymine.

17. The nucleoside base oligomer of Claim 13 wherein at least one B is cytosine.

18. The nucleoside base oligomer of Claim 13, wherein n is at least about 5.

19. A nucleoside monomer of the following formula (X):



(X)

wherein

R^1 is hydrogen, benzyl, -CH_2-p-C_6H_4OH, -CH_2-indol-3-yl, -$CH_2CH_2CH_2CH_2NH_2$, -$CH_2CH_2CH_2NHC(NH)NH_2$, -CH_2-imidazol-4-yl, -CH_2COOH, -$CH_2COO(C_{1-4}$ alkyl), -CH_2CH_2COOH, -$CH_2CH_2COO(C_{1-4}$ alkyl), -CH_2CONH_2, -$CH_2CH_2CONH_2$, -CH_2SH, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkenyl, C_{2-8} alkynyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl or di-(C_{1-4} alkyl)-substituted amino;

R^2 is an amino protecting group;

R^3 is hydrogen, benzyl, -CH_2-p-C_6H_4OH, -CH_2-indol-3-yl, -$CH_2CH_2CH_2CH_2NH_2$, -$CH_2CH_2CH_2NHC(NH)NH_2$, -CH_2-imidazol-4-yl, -CH_2COOH, -$CH_2COO(C_{1-4}$ alkyl), -CH_2CH_2COOH, -$CH_2CH_2COO(C_{1-4}$ alkyl), -CH_2CONH_2, -$CH_2CH_2CONH_2$, -CH_2SH, $CH_2CH_2SCH_3$, C_{1-12} alkyl, C_{2-8} alkenyl, C_{2-8} alkynyl, C_{5-8} cycloalkyl, aryl, heteroaryl, or aryl or heteroaryl which is mono, di, or

trisubstituted independently with halogen, nitro, C_{1-4} alkyl, C_{1-4} alkoxy, trifluoromethyl.

R^4 is hydrogen or a carboxylic acid protecting group;

B is a purine or pyrimidine nucleoside base provided that if R^1 and R^3 are hydrogen, B is a purine nucleoside base,
or an acid or base addition salt thereof.

20. The monomer of Claim 19, wherein B is a purine nucleoside base.

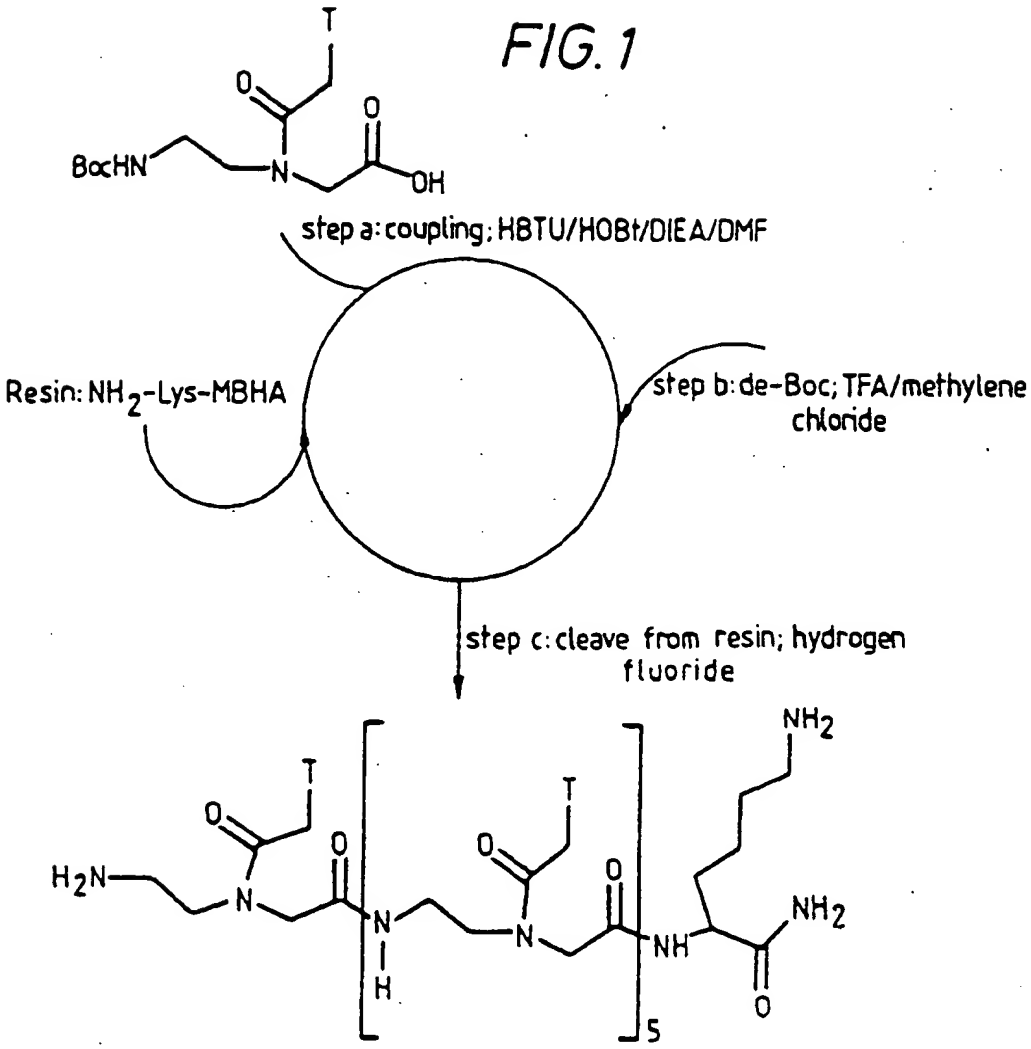
21. A method of affecting genetic material which comprises administering to the genetic material a compound as claimed in any one of claims 13 to 18.

22. A pharmaceutical formulation comprising a compound as claimed in any one of claims 1 to 8 or claims 13 to 18 together with a pharmaceutically acceptable carrier therefor.

23. A compound as claimed in any one of claims 1 to 8 or claims 13 to 18 for use in medicine.

24. A compound as claimed in any one of claims 1 to 8 or claims 13 to 18 for use in the manufacture of a medicament for the treatment of a condition which may be ameliorated by affecting genetic material.

FIG. 1


T
O
O
BocHN
N
OH
step a: coupling; HBTU/HOBt/DIEA/DMF
Resin: NH2-Lys-MBHA
step b: de-Boc; TFA/methylene chloride
step c: cleave from resin; hydrogen fluoride
NH2
H2N
NH
H
S

(Ib)

FIG. 2

Analysis of RNA binding potential by test compound

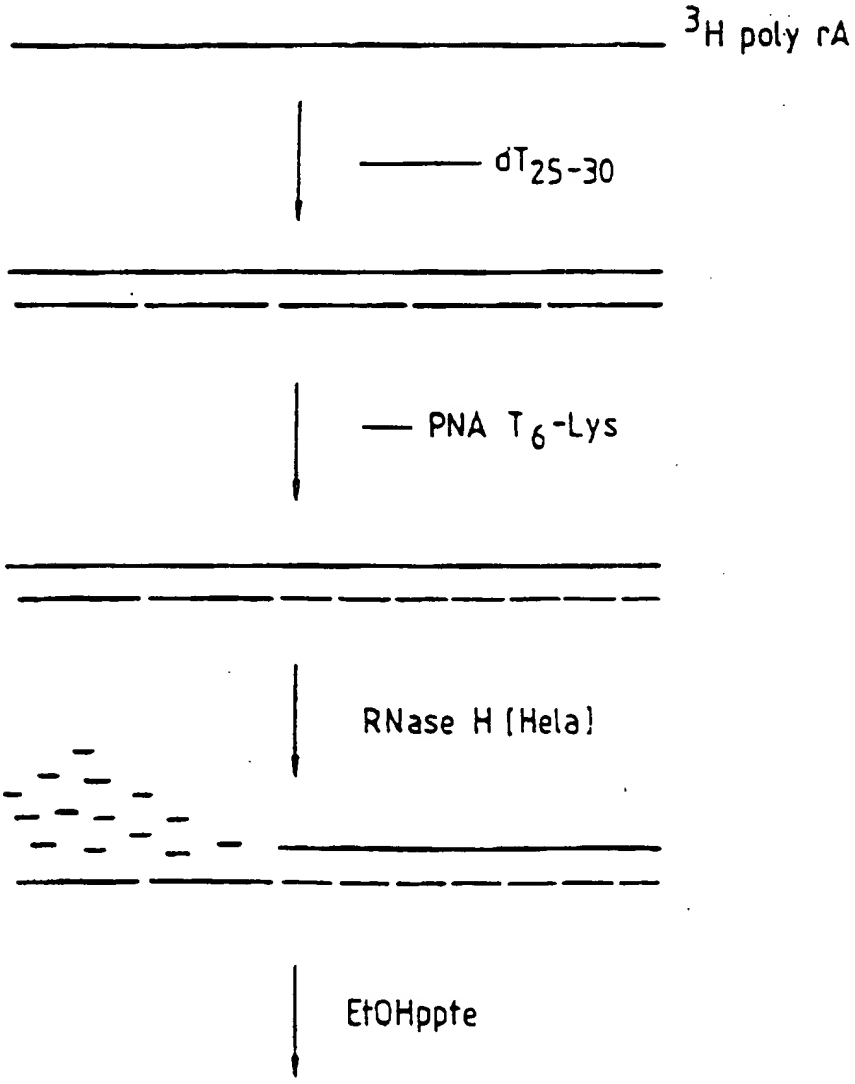
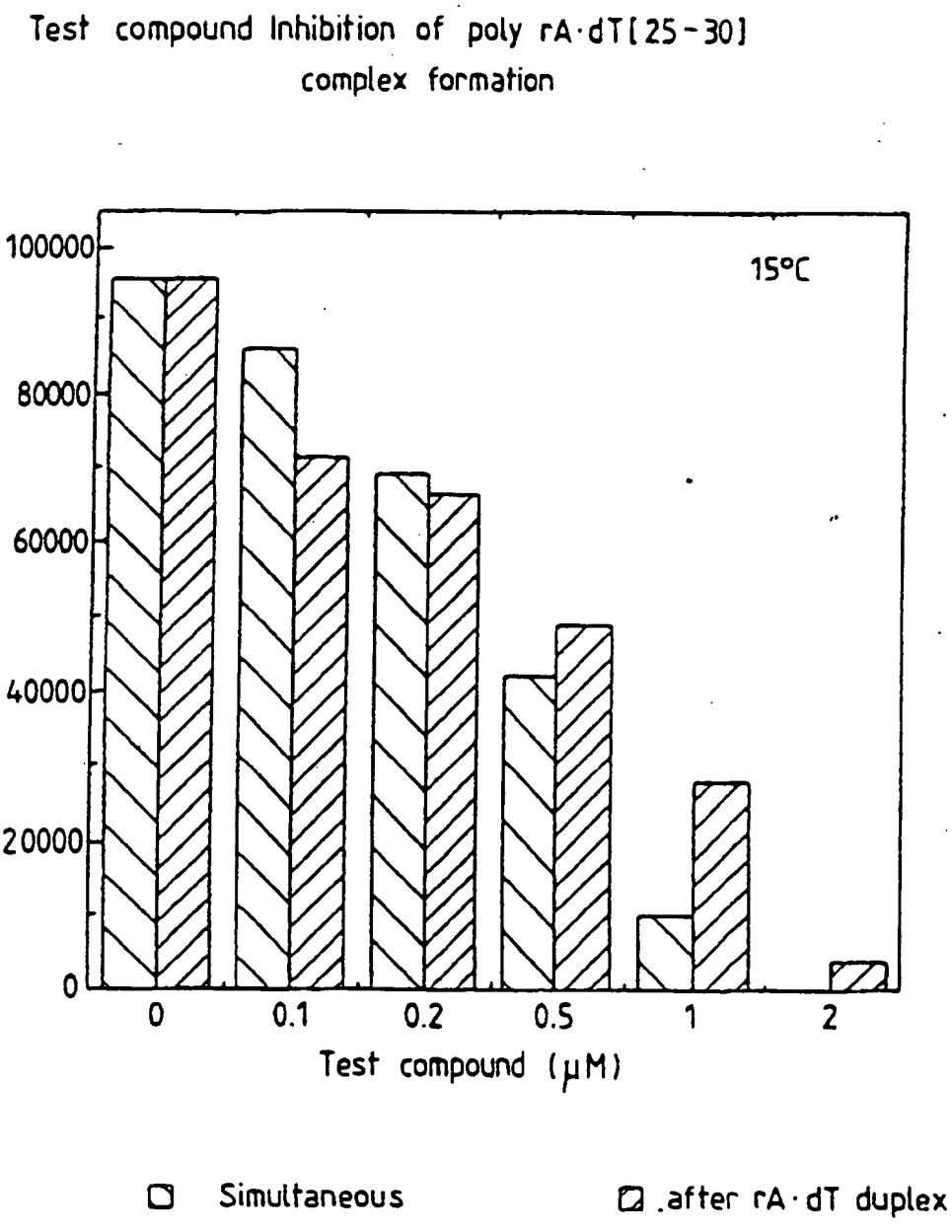


FIG. 3


Test compound Inhibition of poly rA·dT[25-30] complex formation
15°C
100000
80000
60000
40000
20000
0
0
0.1
0.2
0.5
1
2
Test compound (μM)
Simultaneous
after rA·dT duplex

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/10921

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C07H 19/06, 19/16; C08L 77/00; A61K 31/785; [redacted], 31/505, 31/52
US CL :544/264, 242; 514/269, 261
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 544/264, 242; 514/269, 261

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS AND CAS ONLINE: peptide, oligonucleosides, internucleoside linkages, polyamide

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Journal of Organic Chemistry, Volume 56, issued 1991, Weller et al., "Molecular Modeling of Acyclic Polyamide Oligonucleotide Analogues", pages 6000 - 6006, esp. Chart I and Figure 1.	1 - 24
A	Journal of Organic Chemistry, Volume 56, issued 1991, Huang et al., "Acyclic Nucleic Acid Analogues: Synthesis and Oligomerization of gamma,4-Diamino-2-oxo1(2H)-pyrimidinepentanoic Acid and delta-4-Diamino-2-oxo-1(2H)-pyrimidinehexanoic Acid", pages 6007 - 6018, esp. page 6009, Scheme VII.	1 - 24

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 25 JANUARY 1993	Date of mailing of the international search report 23 FEB 1993
Name and mailing address of the ISA/ Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer GARY L. KUNZ Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/10921

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Science, Volume 254, issued 06 December 1991, Nielsen et al., "Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide", pages 1497 - 1500, esp. Figure 1 and page 1498, column 1, last two lines to column 2, lines 1 - 5.	1 - 24